THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[SEVENTH SERIES.]

SUPPLEMENT, FEBRUARY 1935.

XXVI. *Catalytic Combustion at High Temperatures.*

By W. DAVIES, *B.Sc., Ph.D., A.M.I.Mech.E.**

[Plate XIII.]

THE catalytic combustion of carbon monoxide and hydrogen on the surfaces of platinum and palladium wires has been investigated over a range of temperatures extending to nearly the melting-points of the metals by a method which has been described in a previous paper †.

In the course of his well-known investigations of the mechanism of the reactions $2H_2+O_2=2H_2O$ and $2CO+O_2=2CO_2$ on the surface of platinum, Langmuir ‡ measured the rate of combustion of these gases on an electrically heated platinum filament at exceedingly low densities, by observing the rate of fall of the pressure in a closed reaction vessel. In an experiment with carbon monoxide he found that when the filament was raised to a white heat the rate of combustion was reduced to about one-sixth of that obtained at a dull red heat, and he inferred that at the higher temperature, reaction only took place on the cooled ends of the filament.

* Communicated by Prof. W. T. David, Sc.D., M.Inst.C.E.

† Davies, "The Rate of Heating of Wires by Surface Combustion," Phil. Mag. xvii. p. 233 (1934).

‡ Trans. Far. Soc. xvii. p. 624 (1922).

The experiments described in this paper confirm this result. They show that the catalytic combustion of carbon monoxide is entirely suppressed on platinum wires, and also on palladium wires, at a temperature of about 1200° C.; but it has been found, however, that this interruption is only of a temporary character, since the wires recover their activity at higher temperatures and catalytic combustion then proceeds, though at a diminishing rate, up to nearly the melting-points of the wires.

This discontinuity in the catalytic action of the wires is not observed in experiments with hydrogen, except when the wires have been used previously in carbon monoxide mixtures, or have been exposed to the atmosphere for a long period. Under these circumstances a partial or complete interruption of catalytic combustion occurs at about 1200° C. when the wires are first heated up in a hydrogen mixture, but it does not occur again, as with carbon monoxide, when the experiment is repeated.

These experiments appear to indicate that the catalytic surface of platinum and palladium is of a composite nature, as suggested by Taylor *, and an attempt is made to interpret the results on the basis of Taylor's theory of the solid catalytic surface, with the aid of Langmuir's theory of the mechanism of the catalytic oxidation of carbon monoxide and hydrogen on platinum.

The experiments show, incidentally, that a wire of ·001 inch in diameter can be heated to a temperature of at least 1250° C. in the most explosive mixture of carbon monoxide and air without causing ignition, and without any indication that combustion in the gas-phase is taking place on the surface of the wire.

The experimental method is based on the direct measurement of the rate of heating of a metallic wire by surface combustion when it is suspended freely in a combustible mixture contained in a closed vessel. The wire is partially heated by an electric current and the rate of heating by combustion is found by comparing the electrical energy required to produce a given rise of temperature in the wire, when surface combustion is taking place, with that required to produce the same rise of temperature when combustion is suppressed by eliminating either the combustible gas or the oxygen.

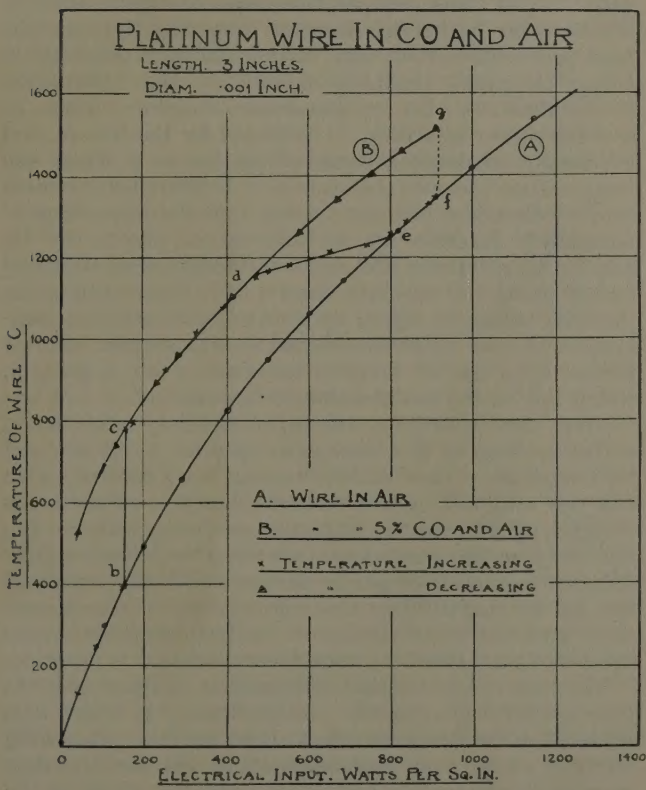
* Proc. Roy. Soc. A, cviii. p. 105 (1925).

Results of Experiments.

(1) *Platinum Wire in Carbon Monoxide and Air.*

The curve A in fig. 1 shows the relationship between the rise of temperature and the input of electrical energy,

Fig. 1.



in watts per sq. in. of surface in a platinum wire ·001 inch in diameter and 3 inches long when the wire is suspended horizontally in air.

When carbon monoxide is added to the air the temperature of the wire rises in practically the same manner as in air up to about 400° C., but at this point catalytic

combustion starts abruptly on its surface, and its temperature is thereby raised almost instantaneously to a higher value, as indicated by the vertical line *bc*. The wire is now heated by combustion as well as by the current, and its temperature rises with increase of electrical energy as shown by the line *cd*, until it reaches about 1160° C. At this temperature the catalytic activity of the wire begins to decrease very rapidly, and the heat produced in it by any further addition of electrical energy is nearly counterbalanced by the retardation of combustion. Its temperature, therefore, rises at a much slower rate now, as indicated by the line *de*, and eventually combustion ceases altogether at *e*, where the temperature becomes the same as it is in air for the same rate of electrical heating. After this the wire remains completely inactive for an interval of about 100° C., and its temperature rises with the increase of electrical energy along the air-curve from *e* to *f*, but at this point the wire begins to regain its activity and catalytic combustion is once more established on its surface, accompanied by a rise of temperature from *f* to a point *g*, which lies on the continuation of the curve *cd*.

After this transition the wire remains catalytically active as long as the combustion process continues, and its temperature rises and falls along the curve *cdg* with the rise and fall of the current; but if combustion is stopped, by reducing the current to zero, or by withdrawing the wire from the gaseous mixture when the latter contains sufficient combustible gas to sustain combustion without the aid of the current, the wire reverts to its original state and the interruption of catalytic activity occurs again at about 1200° C., when the experiment is repeated.

The recurrence of surface combustion at about 1360° C. does not take place in the abrupt manner in which it is initiated in the first instance at about 400° C. It usually starts at one or more isolated points on the wire and then spreads slowly from these points until the whole of the wire is glowing uniformly.

The series of photographs which are reproduced in Pl. XIII. show how the wire appears at different stages during this period. The first photograph, at A, was taken when the wire was in the state represented by the point *f* in fig. 1, when no catalytic combustion was taking place, and the wire was, therefore, heated

by the current alone to a temperature of about 1360° C. The current was kept constant during this experiment, and shortly afterwards combustion was initiated at a point near the right-hand side of the wire, from which it spread, as shown by the extension of the region of greater luminosity in the lower photographs. These photographs were taken at equal intervals of about one minute until combustion had spread over practically the whole length of the wire as shown at E. The more highly luminous region of the wire, shown in the lower photographs, was heated by combustion, as well as by the current, to a temperature of about 1560° C., whilst the remainder of the wire was heated by the current alone to a temperature of about 1360° C.

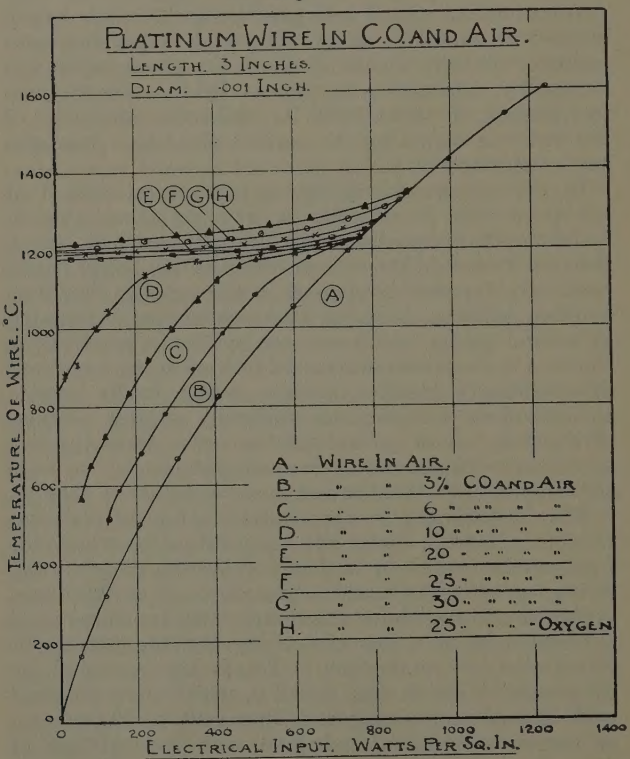
In this particular experiment combustion started at one point only, from which it travelled at an approximately uniform speed towards the left, and had nearly reached the end of the wire in this direction before it had made any appreciable progress in the opposite direction. It often happens, however, that combustion is initiated at several points simultaneously, and its progress along the wire is sometimes obstructed by very minute portions of abnormally inactive surface, which finally remain as cold spots bridging the uniformly glowing sections of the wire. As a general rule, however, these inactive spots can be eliminated by increasing the current to raise the temperature of the wire at these places above 1360° C.

The discontinuity in the catalytic action of platinum at about 1200° C. limits the temperature to which the wire can be heated by surface combustion alone in any mixture of carbon monoxide and air, because any tendency of the wire to rise above this temperature is counteracted automatically by a loss of activity and the consequent retardation of combustion. This is demonstrated by the results shown in figs. 2 and 3, which were obtained with mixtures containing from 3 per cent. to 30 per cent. of carbon monoxide in air and with one mixture of 25 per cent. of carbon monoxide and oxygen. The curves B to H in fig. 2 show the temperatures attained by the wire in these mixtures with different inputs of electrical energy after surface combustion had been initiated at about 400° C., and the curve in fig. 3 shows the rise of temperature produced by surface combustion in the various mixtures when the input of electrical energy was

200 watts per sq. in., before the wire had been made active at temperatures above 1200° C.

In mixtures containing more than about 15 per cent. of carbon monoxide the temperature of the wire is immediately raised to the limit of about 1200° C. as soon as

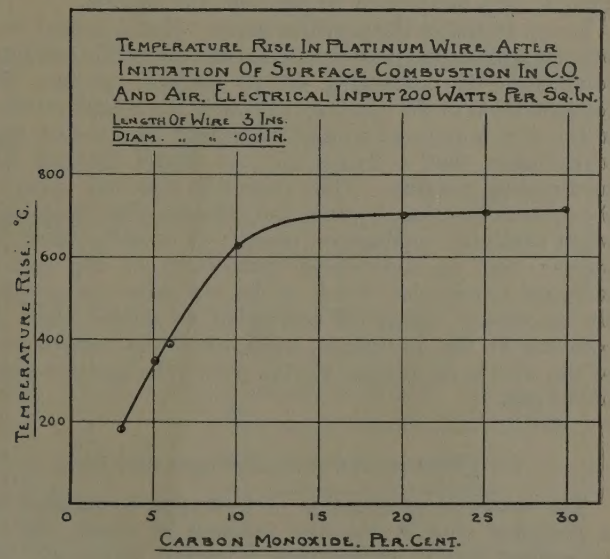
Fig. 2.


PLATINUM WIRE IN CO AND AIR.
LENGTH. 3 INCHES.
DIAM. ·001 INCH.
TEMPERATURE OF WIRE. °C.
ELECTRICAL INPUT. WATTS PER SQ. IN.
A. WIRE IN AIR.
B. " " 3% CO AND AIR
C. " " 6 " " " "
D " " 10 " " " "
E. " " 20 " " " "
F. " " 25 " " " "
G. " " 30 " " " "
H. " " 25 " " " OXYGEN

surface combustion begins, after which the temperature remains practically constant and independent of the current over a certain range of electrical input, since any tendency towards a further rise of temperature, whether it is due to the current or to combustion, is counteracted by a corresponding loss of catalytic activity.

The combined heating effects of the current and surface combustion during this period is thus practically constant and the curves E to H are nearly horizontal, but this balance can, of course, only be maintained until these curves reach the air-curve A, since combustion then ceases altogether, after which the temperature of the wire must rise along the air-curve with any further increase of electrical energy Almost as soon as this occurs, however, the wire regains its catalytic activity in the manner which

Fig. 3.



has been described previously, and it is instantaneously raised to the melting-point at the first spot where surface combustion is established for the second time, with the result that ignition in the gas-phase takes place and the gaseous mixture explodes.

This has been found to occur sometimes before the curves E to H in fig. 2 actually reach the air-curve A. The temperatures indicated by these curves also vary in an apparently accidental manner with different wires between about 1180° C. and 1250° C., probably on account

of the uneven quality of the catalytic surface at this critical juncture.

The ignition temperature of an explosive mixture of carbon monoxide and air is of the order of 600° C., but these experiments show that it is possible to heat a wire of ·001 inch diameter to at least 1300° C. in such a mixture without causing ignition, and even without any indication that the wire is heated to a measurable extent by combustion in the gas-phase near its surface. Thus, at the junction of the curve G and A in fig. 2, the wire is obviously heated by the current alone to a temperature of just over 1300° C. in a mixture of 30 per cent. CO and air.

It was stated in the previous paper * that a heated wire will not ignite an explosive gaseous mixture while catalytic combustion is proceeding on its surface, because the concentration of the combustible gas in the neighbourhood of the wire is reduced to such an extent by the slow rate of diffusion that a flame will not travel through the surrounding gas-film. This, evidently, does not apply in the present instance, because the diffusion-film disappears when catalytic combustion ceases; it would, therefore, appear that no appreciable proportion of the carbon monoxide molecules which strike the wire can acquire the necessary energy of activation to enable them to combine in the gas-phase, even when the temperature of the wire is more than 600° C. above the ignition-point of the gas.

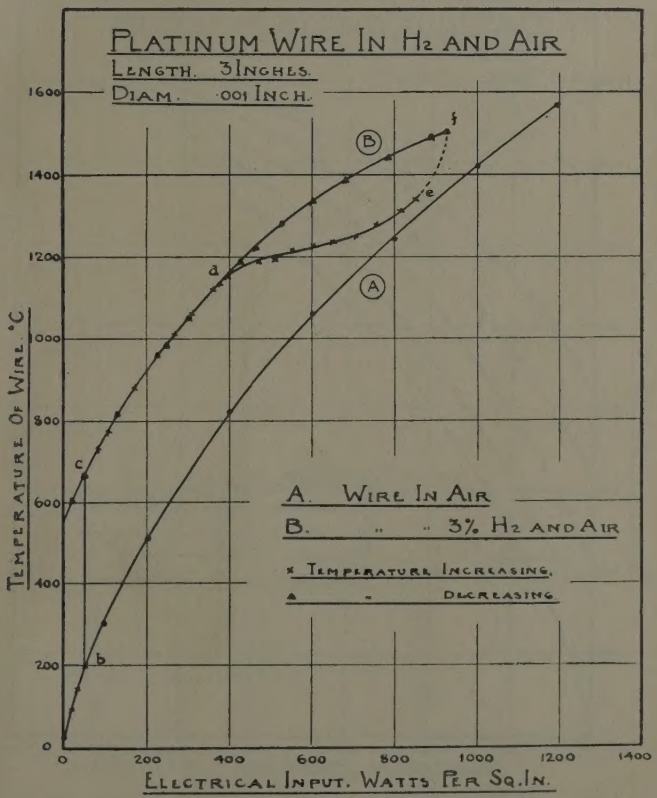
(2) *Platinum Wires in Hydrogen and Air.*

The catalytic combustion of hydrogen is initiated on a platinum wire in the first instance at about 200° C., as against 400° C. in the case of carbon monoxide, but no interruption of activity occurs at 1200° C. in experiments with hydrogen unless the wire has been used previously with carbon monoxide, or has been exposed to the air for a long period. Under these circumstances catalytic combustion is partially, and sometimes completely, suppressed between 1200° C. and 1300° C. when the wire is heated up for the first time in a hydrogen mixture, but it does not occur again if the experiment is repeated within a few hours.

* Phil. Mag. xvii. p. 250 (1934).

The results shown in fig. 4 were obtained with a wire which had been used immediately before in an experiment with a mixture of 8 per cent. of carbon monoxide and air. During the first time of heating in the hydrogen experiment the temperature of the wire increased as shown by the

Fig. 4.

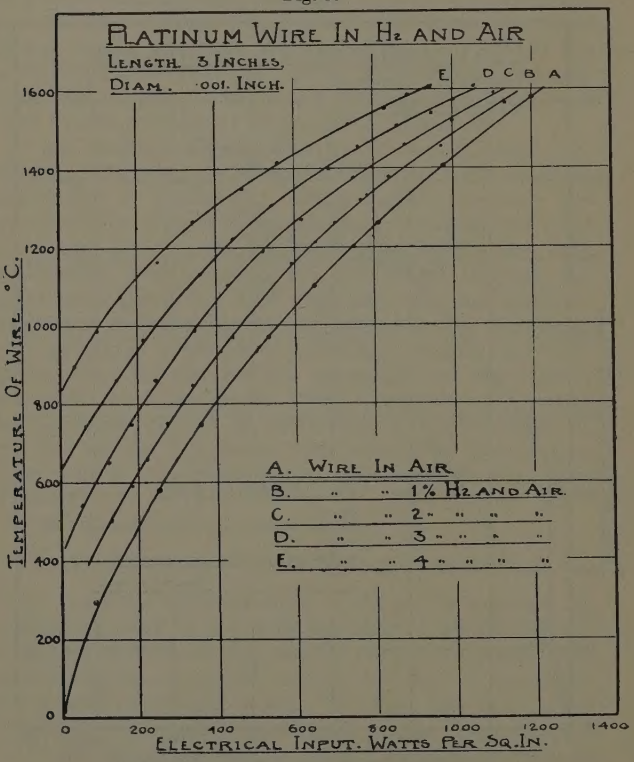


curve *b c d e f*, but in all subsequent experiments with the same mixture it followed the curve *b c d f*. The turning-point in the curve at *e* during the first heating lies to the left of the air-curve, therefore the catalytic action of the wire was only partially interrupted in this particular experiment, and this is typical of the results usually

obtained under the circumstances which have been described.

The interruption of the activity of the wire during the first heating in a hydrogen experiment appears to become less in extent as the percentage of hydrogen in the air

Fig. 5.

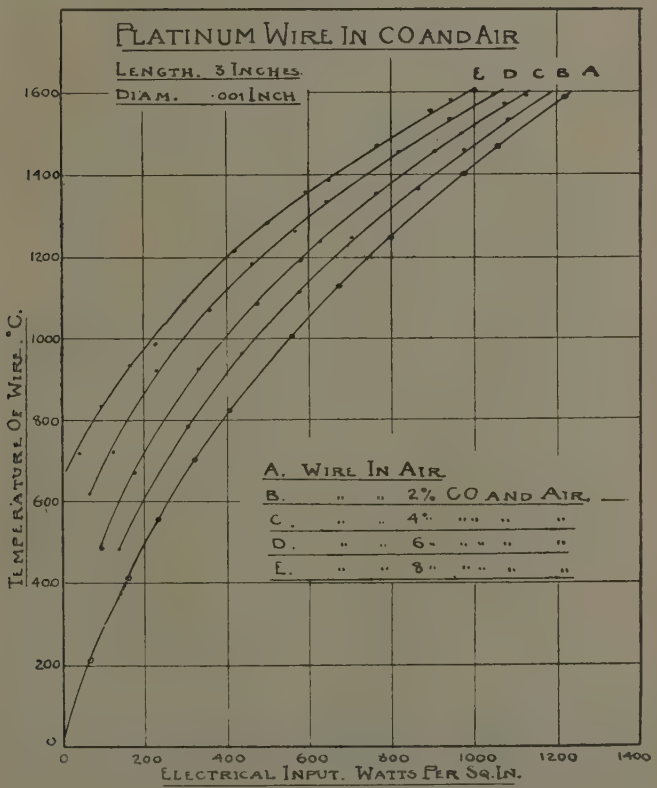


is increased; therefore it is not possible to obtain results similar to those shown in fig. 2 with hydrogen mixtures, since the initial rise of temperature is not restricted to 1200° C., and ignition, followed by a partial explosion, generally occurs immediately after surface combustion starts on the wire in mixtures containing even less than 10 per cent. of hydrogen.

(3) *Reaction Rates at High Temperatures.*

The results shown in fig. 5 were obtained with a continuously active wire in mixtures of hydrogen and air containing from 1 per cent. to 4 per cent. of hydrogen,

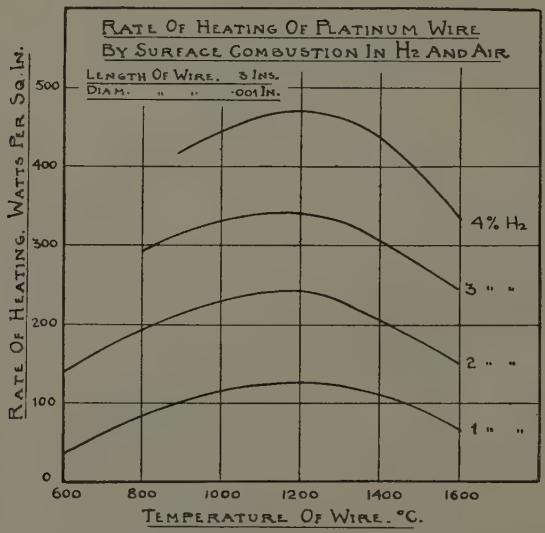
Fig. 6.



and those shown in fig. 6 were obtained with from 2 per cent. to 8 per cent. of carbon monoxide and air, by first heating the wire rapidly to about 1500° C., to make it catalytically active at high temperatures, and then reducing the current to obtain the readings at lower temperatures.

The reaction rates at different temperatures in these experiments, expressed in terms of the rate of heating of the wire, are shown in figs. 7 and 8. As long as the wire promotes the reaction at a sufficient rate to use up all the combustible gas as quickly as it reaches the surface, the rate of heating increases with the temperature of the wire on account of the acceleration of diffusion. This condition appears to hold up to about 1200° C.;

Fig. 7.



then the rate of heating begins to decrease, slowly at first, but more rapidly at temperatures above 1400° C.

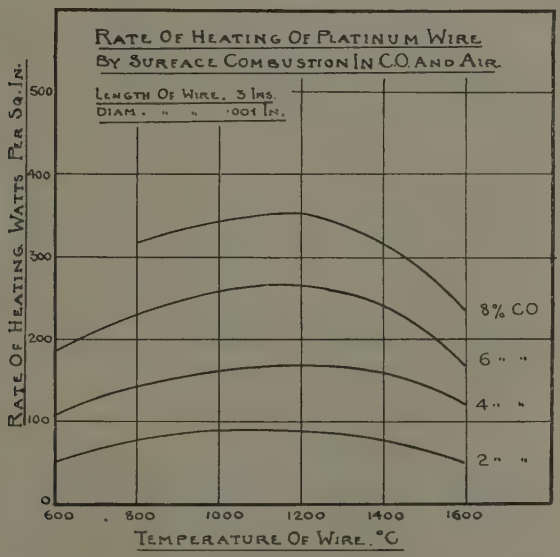
The decrease in the rate of heating at high temperatures is probably due to the elimination of the active surface, but diffusion still appears to be the controlling factor, since the ratio of the rates of heating with equal amounts of hydrogen and carbon monoxide remains more or less the same throughout at about 2·7, whereas if the transfer of the combustible gas to the wire were not restricted by diffusion, the rate of heating would be greater with carbon monoxide than with hydrogen, because the heat

of combustion per gram molecule of carbon monoxide is greater than that of hydrogen.

(4) *Experiments with Palladium Wires.*

The results obtained with palladium wires are, in general, very similar to those obtained with platinum wires. In experiments with carbon monoxide, catalytic combustion

Fig. 8.



starts in the first instance at about 400° C., as in the case of platinum, but it does not usually spread over the whole of the wire until the temperature has been raised to about 1000° C. *. The wire then glows uniformly, and the rate of heating at this temperature, and also at lower temperatures, is the same as that obtained with the platinum wire. When the temperature is increased still further catalytic combustion ceases at about the same point as in the case of platinum, *i. e.*, at about 1200°C., and it begins again at about 1350° C. to 1400° C., in the same way as that shown in the photographs in Pl XIII.

* Phil. Mag. xvii. p. 243 (1934).

Since, however, the melting-point of palladium is now being approached, considerable difficulty is experienced in bringing the wire into an active state over the whole of its length without overheating it and causing it to fuse at one or more points; but, when this has been successfully accomplished, the results, up to a temperature of about 1420° C., are, in every respect, similar to those obtained with a platinum wire.

No detailed reference need be made to the experiments with palladium wires in hydrogen mixtures, since the results are the same as those obtained with platinum wires in all respects, except that catalytic combustion starts in the first instance at about 120° C., as against 200° C. in the case of platinum.

It should be pointed out that these initial temperatures are those which are observed when smooth wires with polished surfaces are used. In the course of time the surface of the wire gradually changes from a smooth to a rough condition, and turns into a grey colour. This change is invariably accompanied by a lowering of the temperature at which catalytic combustion is first initiated in experiments with hydrogen and with carbon monoxide. Detailed observations at high temperatures with a wire in this condition have not yet been made.

Discussion.

A full discussion of the results described in this paper is reserved until further experiments with other gases are completed, but it may be useful, at this stage, to indicate briefly how a possible interpretation is suggested by Langmuir's theory of the catalytic oxidation of carbon monoxide and hydrogen on platinum and Taylor's theory of the solid catalytic surface.

Langmuir * postulates a heterogeneous phase on the platinum surface which he considers, on theoretical grounds and from experimental evidence, to be restricted to a film of monomolecular thickness. The gaseous constituent of this film may be resolved into its atomic form, or it may be oriented in a manner which renders it either more reactive or less reactive than when it is in the gas-phase.

* Trans. Far. Soc. xvii. p. 621 (1922).

When the atmosphere above the platinum surface contains several gases, selective adsorption generally takes place according to the temperature of the wire. Thus, in a mixture of carbon monoxide and oxygen, Langmuir has shown that carbon monoxide only is adsorbed until the temperature of the platinum is raised to about 400° C. At this point the carbon monoxide evaporates and is replaced by oxygen, and since carbon monoxide in the heterogeneous phase on platinum is less reactive than when it is in the gas-phase, and oxygen, on the other hand, becomes exceedingly active in the process of adsorption, the transition which occurs at about 400° C. is immediately followed by the rapid oxidation of all the carbon monoxide molecules which strike the wire, so that the initiation of surface combustion at this temperature in the experiments with carbon monoxide is readily explained.

Langmuir predicted that the platinum filament, which he used in his low-pressure experiments, would lose its activity at a white heat because the oxygen would eventually be driven from the surface, and the wire, thus freed from all adsorbed gas, would cease to act as a catalyst. The experiments described in this paper show that, whilst the catalytic activity of the wire is interrupted at about 1200° C. in the manner predicted by Langmuir, the interruption is only temporary, since the wire regains its activity after it has been heated by the current to a higher temperature.

According to Taylor *, a solid catalytic surface consists of irregularly distributed active centres, which constitute only a small fraction of the total area of the surface, and, from a consideration of differential catalytic phenomena and the disproportionate effects produced by changes in the apparent surface of a catalyst when it is subdivided, he concludes that the adsorptive properties of different centres vary according to their structural configuration.

In the light of this theory, the two stages of catalytic activity exhibited by the wire in the experiments with carbon monoxide point to the existence of two principal groups of active centres, on one of which carbon monoxide is adsorbed up to a temperature of about 400° C. and on the other up to about 1360° C. to 1400° C. The interruption of catalytic combustion at about 1200° C., when the

* Proc. Roy. Soc. A, cviii. p. 105 (1925).

temperature is rising, may be explained on the assumption that oxygen evaporates from the first group of centres before the second group becomes available, and the continuity of combustion at this point when the temperature is falling is probably due to the protection of the active centres against the poisoning action of carbon monoxide by the diffusion film †, which limits the rate at which the carbon monoxide reaches the wire to such an extent that it cannot compete with the oxygen for the empty spaces left vacant after reaction.

The decrease in the rate of heating at very high temperatures is probably caused by the progressive elimination of the active centres in the second group due to the evaporation of oxygen. Although an increasing proportion of the carbon monoxide molecules which strike the wire now do not react, the ratio of the rate of heating with carbon monoxide and with hydrogen remains more or less the same, as shown by the curves in figs. 7 and 8; therefore diffusion is still the controlling factor, but it is now confined to the spaces above and around the centres which remain active.

According to Langmuir, the catalytic combustion of hydrogen takes place in similar stages, but the transition temperatures are generally lower. The continuous activity of the wire in the hydrogen experiments might, therefore, be explained on the assumption that hydrogen evaporates from the second group of active centres before the oxygen leaves the first group. Under these circumstances combustion would become continuous, although the catalytic process itself would be discontinuous in the sense that a more or less abrupt change of active centres takes place at about 1200° C.

The results shown in fig. 2 indicate that a wire of ·001 inch diameter can be raised to a temperature of over 1200° C. in explosive mixtures of carbon monoxide and air without any sign of combustion in the gas-phase, since, when the curves E to H meet the air-curve A, it is evident that the wire is then heated by the current alone in these mixtures to a temperature which is more than twice as high as the ignition-point of the gas.

The problem of the ignition of explosive gaseous mixtures by heated wires is now being investigated by a similar method to that used in the present experiments,

* Phil. Mag. xvii. p. 250 (1934).

but it may be pointed out that there is no anomaly in these experimental facts. The mass of gas surrounding the wire is cold, and it is probable that the temperature gradient at the surface of the wire corresponds to a fall of a hundred degrees or more in a distance equal to the mean free path. Under these conditions only an infinitesimal proportion of the molecules which strike the wire can acquire the necessary energy of activation to enable them to combine in the gas-phase, even when the temperature of the wire is far above the ignition-point of the gas, as ordinarily defined. Hinshelwood and Prichard * have previously shown that the temperature gradient in a gas surrounding a heated wire is of the order of several hundred degrees in the first few mean free paths, from a consideration of the kinetics of the catalytic decomposition of nitrous oxide on the surface of an electrically heated gold wire. They found that the reaction was entirely monomolecular, and, therefore, took place only on the surface of the wire when the latter was raised to a temperature of 1000° C., which is considerably higher than the temperature at which the bimolecular reaction proceeds in the gas-phase.

My thanks are due to Professor W. T. David for his continued interest in this work.

Engineering Dept.,
The University, Leeds.

XXVII. *Fourier Expansions in the Classical Theory of the Motion of a Rigid Body under no Forces, and their Application in the Bohr Theory of the Asymmetrical Rotator.* By Enos E. Witmer, *Ph.D., Department of Physics, University of Pennsylvania* †.

PRIOR to the development of the new quantum mechanics, the asymmetrical rotator was treated by the writer ‡ in terms of the Bohr theory §. Although

* Proc. Roy. Soc. A, cviii. p. 211 (1925).

† Communicated by the Author.

‡ E. E. Witmer, Proc. Nat. Acad. xii. p. 602 (1926). This paper contains also a statement of the selection rules, without, however, giving their derivation.

§ *Cf.* also G. Thomsen, *Math. Annalen*, xciv. p. 146 (1925); P. Epstein, *Verh. d. Deutschen Phys. Ges.* xviii. p. 398 (1916); *Phys. Zeitschr.* xx. p. 289 (1919).

the more recent work of Kramers and Ittmann, O. Klein, and Casimir * has resulted in a satisfactory solution of the problem in terms of wave-mechanics, it nevertheless appears to be of interest to present at this time certain results originally found in obtaining selection rules and intensities of the asymmetrical top according to the Bohr theory. A possible justification of this lies in the fact that in the course of the work it was necessary to obtain Fourier expansions of the functions occurring in the classical theory of the motion of a rigid body under no forces ; and such expansions are of interest in the classical dynamics of the problem, independently of their application in quantum theory.

Although solutions of the problem of the force-free motion of a rigid body in terms of elliptic and theta functions are given in the standard treatises on the subject †, it appears desirable, for the purpose in hand, to obtain a solution in a slightly different and somewhat simpler form. In what follows we shall indicate the procedure used.

The notation will be that of Born's ' Atommechanik,' (p. 126 ff.), from which a number of formulæ will be taken. A rigid body has three principal moments of inertia, A_x, A_y, A_z, the axes of which are taken as the x, y, z axes, respectively. These are fixed in the body. (X, Y, Z) represents another coordinate system fixed in space, the origin being the centre of gravity of the body. The relation of these two coordinate systems is defined by Euler's angles : θ, the angle between the z and Z axes ; ϕ, the angle between the x-axis and the line of nodes (intersection of the xy and XY planes) ; ψ, the angle between the X-axis and the line of nodes. The conjugate momenta are p_θ, p_ϕ, p_ψ, respectively.

The general nature of the motion can be seen from the following considerations. If D denotes the total angular momentum whose components along the x, y, z axes are respectively D_x, D_y, D_z, the total energy W, which is equal to T, the kinetic energy, is given by

$$W=T=\tfrac{1}{2}\left[\frac{D_x^2}{A_x}+\frac{D_y^2}{A_y}+\frac{D_z^2}{A_z}\right]. \quad . \quad . \quad . \quad (1)$$

* H. A. Kramers and G. P. Ittmann, *Zs. f. Phys.* liii. p. 553 (1929) : lviii. p. 217 (1929) ; lx. p. 663 (1930). O. Klein, *Zs. f. Phys.* lviii. p. 730 (1929). H. B. G. Casimir, *Zs. f. Phys.* lix. p. 623 (1929).

† E. T. Whittaker, ' Analytical Dynamics,' chap. vi. § 69 ; Klein-Sommerfeld, ' Theorie des Kreisels,' ii. chap. vi. § 77 ; R. Grammel, ' Der Kreisel.'

We have also

$$D^2=D_x^2+D_y^2+D_z^2. \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Regarding D_x, D_y, and D_z as variables, eqs. (1) and (2) are respectively the equations of an ellipsoid and a sphere. For all possible motions these two surfaces intersect in a curve which we shall designate C. This curve represents the path of the end-point of the vector of total angular momentum referred to the axes fixed in the body.

For any given value of D a continuous range of values of W is possible. In this paper we shall assume

$$A_x<A_y<A_z. \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

If W is too small, the ellipsoid (1) will lie entirely within the sphere (2). If, however, $W=\frac{D^2}{2A_z}$, the sphere and the ellipsoid will be in contact at the extremity of the z-axis of the ellipsoid, the remainder of the surface of the ellipsoid being within the sphere. This corresponds to pure rotation about the z-axis of the ellipsoid. If

$$\frac{D^2}{2A_z}<W<\frac{D^2}{2A_y},$$

the curve C will be a closed space curve, whose centre lies on the z-axis. When $W=\frac{D^2}{2A_y}$, the curve C will consist of two intersecting circles whose points of intersection are the extremities of the y-axis of the ellipsoid. If

$$\frac{D^2}{2A_y}<W<\frac{D^2}{2A_x},$$

the curve C is a closed curve whose centre lies on the x-axis. Finally, when $W=\frac{D^2}{2A_x}$, the curve C reduces to a point, namely, the extremity of the x-axis of the ellipsoid. In this case we have pure rotation about the x-axis. For the analytical treatment it is necessary to divide the

motions into two classes, I. and II., which are separated by a third class, III. These classes are :—

$$\text{I.} \quad \frac{D^2}{2A} \leq W < \frac{D^2}{2A_y}, \quad . \quad . \quad . \quad . \quad (4)$$

$$\text{II.} \quad \frac{D^2}{2A_y} < W \leq \frac{D^2}{2A_x}, \quad . \quad . \quad . \quad . \quad (5)$$

$$\text{III.} \quad W = \frac{D^2}{2A_y}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The reason for introducing these classes is that in the analytical treatment use is made of elliptic and Θ-functions. If one attempts to treat classes I. and II. simultaneously, it is found that the modulus κ, defined in eq. (17) below, although less than one for motions of class I., is greater than one for motions of class II. Conversely, if κ had been defined so as to be less than one for motions of class II., then it would be greater than one for motions of class I. If one wishes, however, to make use of elliptic functions, κ must be less than or equal to one ; the latter case corresponds to motions of class III., which will not be treated in detail for reasons to be given below. In what follows we shall treat the motions of class I. The theory for class II. is the same except that the rôle of the x and z axes (fixed in the body) must be interchanged. In case III. the motion is either pure rotation around the y-axis fixed in the body, or an aperiodic motion in which the herpolhode has the form of a spiral. Since the character of the motion is not unique, and depends on initial conditions, and since, furthermore, the Bohr theory cannot be applied to aperiodic motions, it seems doubtful whether the Bohr theory can be applied to motions of class III. This case, therefore, will not be considered in detail.

The Hamiltonian has the following form * :—

$$H = T = \tfrac{1}{2}\left\{ \frac{[p_\theta \sin\theta \cos\phi + (p_\psi - p_\phi \cos\theta)\sin\phi]^2}{A_x \sin^2\theta} + \frac{[p_\theta \sin\theta \sin\phi - (p_\psi - p_\phi \cos\theta)\cos\phi]^2}{A_y \sin^2\theta} + \frac{p_\phi^2}{A_z} \right\}, \quad . \quad (7)$$

* E. E. Witmer, Proc. Nat. Acad. xiii. p. 60 (1927).

A considerable simplification results if we choose axes in which the Z-axis (fixed in space) coincides with the vector, D, of total angular momentum.

Then

$$\left.\begin{aligned} p_\theta &= 0, \\ p_\psi &= \mathrm{D}, \\ p_\phi &= \mathrm{D}\cos\theta. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (8)$$

Using Hamilton's equations, and inserting the values (8), one readily obtains

$$\dot{\theta} = \frac{\partial \mathrm{H}}{\partial p_\theta} = \mathrm{D}a \sin\theta \sin 2\phi, \quad . \quad . \quad . \quad (9\,a)$$

where

$$a = \tfrac{1}{2}\left(\frac{1}{\mathrm{A}_x} - \frac{1}{\mathrm{A}_y}\right).$$

and

$$\dot{\psi} = \frac{\partial \mathrm{H}}{\partial p_\psi} = \beta - \mathrm{D}a\cos 2\phi, \quad . \quad . \quad . \quad (9\,b)$$

where

$$\beta = \frac{\mathrm{D}}{2}\left(\frac{1}{\mathrm{A}_x} + \frac{1}{\mathrm{A}_y}\right).$$

Further

$$\dot{\phi} = \frac{\partial \mathrm{H}}{\partial p_\phi} = -\mathrm{D}\cos\theta\left(\frac{\sin^2\phi}{\mathrm{A}_x} + \frac{\cos^2\phi}{\mathrm{A}_y}\right) + \frac{\mathrm{D}\cos\theta}{\mathrm{A}_z}$$

$$= \mathrm{D}\cos\theta\left[\frac{1}{\mathrm{A}_z} - \left(\frac{\sin^2\phi}{\mathrm{A}_x} + \frac{\cos^2\phi}{\mathrm{A}_y}\right)\right] \quad . \quad . \quad (10)$$

Now $\cos^2\theta$ may be expressed in terms of ϕ, as follows (Born, *l. c.* p. 132) :—

$$\cos^2\theta = \frac{\dfrac{2\mathrm{W}}{\mathrm{D}} - \left(\dfrac{\sin^2\phi}{\mathrm{A}_x} + \dfrac{\cos^2\phi}{\mathrm{A}_y}\right)}{\dfrac{1}{\mathrm{A}_z} - \left(\dfrac{\sin^2\phi}{\mathrm{A}_x} + \dfrac{\cos^2\phi}{\mathrm{A}_y}\right)} \quad . \quad . \quad . \quad . \quad (11)$$

Eq. (10), therefore, becomes

$$\dot{\phi} = \sqrt{\left[2\mathrm{W} - \mathrm{D}^2\left(\frac{\sin^2\phi}{\mathrm{A}_x} + \frac{\cos^2\phi}{\mathrm{A}_y}\right)\right]\left[\frac{1}{\mathrm{A}_z} - \left(\frac{\sin^2\phi}{\mathrm{A}_x} + \frac{\cos^2\phi}{\mathrm{A}_y}\right)\right]}. \quad . \quad . \quad . \quad (12)$$

Let us now introduce f, the distance from the centre of gravity to the invariable plane. It is given by the relation

$$f^2 = \frac{2W}{D^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

Also let

$$\left.\begin{aligned} j^2 &= \frac{1}{A_x}, \\ k^2 &= \frac{1}{A_y}, \\ l^2 &= \frac{1}{A_z}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

In this notation (3) becomes

$$j^2 > k^2 > l^2, \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

and eqs. (4), (5), and (6) assume the form

$$\text{I.} \quad k^2 > f^2 \geq l^2, \quad . \quad . \quad . \quad . \quad (16\,a)$$

$$\text{II.} \quad j^2 \geq f^2 > k^2, \quad . \quad . \quad . \quad . \quad (16\,b)$$

$$\text{III.} \quad k^2 = f^2 . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16\,c)$$

Now let

$$\left.\begin{aligned} m &= -\frac{j^2-k^2}{j^2-f^2}, \\ \kappa^2 &= \frac{(f^2-l^2)(j^2-k^2)}{(k^2-l^2)(j^2-f^2)} = -\kappa_1{}^2 m. \end{aligned}\right\} \quad . \quad . \quad (17)$$

Then

$$\kappa^2 < -m < 1. \quad . \quad . \quad . \quad . \quad . \quad (17\,a)$$

Equation (12) now becomes

$$\dot{\phi} = D\sqrt{[f^2-(j^2\sin^2\phi+k^2\cos^2\phi)][l^2-(j^2\sin^2\phi+k^2\cos^2\phi)]}. \quad . \quad . \quad . \quad (18)$$

In what follows it will be convenient to make use of the transformation, given by Epstein *,

$$\cos^2\phi = \frac{\cos^2\chi}{1+m\sin^2\chi} \quad . \quad . \quad . \quad . \quad (19)$$

Introducing the new variable, χ, into (18), it is found,

* P. Epstein, *loc. cit.*

after some algebraic manipulation, that eq. (18) reduces to the form

$$\frac{\mathrm{D}\alpha t}{\sqrt{m+1}} \equiv \tau = \int_0^{\chi} \frac{d\chi}{\sqrt{1-\kappa^2 \sin^2 \chi}}, \quad . \quad . \quad . \quad (20)$$

where

$$\alpha = \sqrt{(f^2-k^2)(l^2-k^2)}, \quad . \quad . \quad . \quad . \quad (20\,a)$$

Hence, in terms of Jacobian elliptic functions,

$$\left.\begin{aligned} \chi &= \mathrm{am}\ \tau, \\ \sin\chi &= \mathrm{sn}\ \tau, \\ \cos\chi &= \mathrm{cn}\ \tau. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (21)$$

We can now express $\cos\theta$ and $\sin\theta$ as functions of χ, and hence the time, by means of eq. (11), which, in terms of χ, assumes the following fairly simple form:

$$\cos\theta = \sqrt{\frac{m+\kappa^2}{m}} \frac{1}{\sqrt{1-\kappa^2 \sin^2\chi}}$$

$$= \sqrt{\frac{m+\kappa^2}{m}} \frac{1}{\sqrt{1-\kappa^2 \mathrm{sn}^2\tau}} = \sqrt{\frac{m+\kappa^2}{m}} \frac{1}{\mathrm{dn}\,\tau}. \quad . \quad (22)$$

We also obtain from this

$$\sin\theta = \sqrt{-\frac{\kappa^2}{m}} \sqrt{\frac{1+m\sin^2\chi}{1-\kappa^2\sin^2\chi}} = \sqrt{-\frac{\kappa^2}{m}} \sqrt{\frac{1+m\,\mathrm{sn}^2\tau}{1-\kappa^2\mathrm{sn}^2\tau}}. \quad . \quad . \quad . \quad (23)$$

It now remains to determine ψ as a function of the time. We have (eq. 9 *b*)

$$\dot{\psi} = \beta - \mathrm{D}\alpha \cos 2\phi.$$

Introducing the variable χ, this becomes

$$\dot{\psi} = \beta - \mathrm{D}\alpha\left[1 - \frac{2(m+1)\sin^2\chi}{1+m\sin^2\chi}\right]. \quad . \quad . \quad . \quad (24)$$

Integrating, and remembering that

$$dt = \frac{\sqrt{m+1}}{\mathrm{D}\alpha} d\tau,$$

$$\psi = (\beta - \mathrm{D}\alpha)t + \frac{2(m+1)^{3/2}\alpha}{\alpha} \int_0^{\tau} \frac{\mathrm{sn}^2\,\tau}{1+m\,\mathrm{sn}^2\,\tau} d\tau. \quad . \quad (25)$$

Let us now introduce the variable μ by the relation

$$m=-\kappa^2 \operatorname{sn}^2 \mu. \quad . \quad . \quad . \quad . \quad . \quad (26)$$

It is then found that

$$\kappa^2 \operatorname{sn} \mu \operatorname{cn} \mu \operatorname{dn} \mu=\sqrt{(m+1)(\kappa^2+m)(-m)}, \quad . \quad (27)$$

and hence the equation (25) can be written *

$$\psi=(\beta-\mathrm{D}\alpha)t+\frac{2(m+1)\alpha}{\alpha\sqrt{(\kappa^2+m)(-m)}}\Pi(\tau, \mu), \quad . \quad (28)$$

where $\Pi(\tau, \mu)$ is the elliptic integral of the third kind. Further, it can easily be shown that

$$\frac{2(m+1)\alpha}{\alpha\sqrt{(\kappa^2+m)(-m)}}=-\sqrt{-1}=-i. \quad . \quad . \quad (29)$$

Hence we have

$$\psi=(\beta-\mathrm{D}\alpha)t-i\Pi(\tau, \mu). \quad . \quad . \quad . \quad . \quad (30)$$

A more convenient form can be obtained in terms of Θ-functions. Making use of the expression † for the elliptic integral of the third kind in terms of Θ-functions, eq. (30) becomes

$$\psi=(\beta-\mathrm{D}\alpha)t-i\tau\frac{\Theta'(\mu)}{\Theta(\mu)}+\tfrac{1}{2}\,i\log\frac{\Theta(\tau+\mu)}{\Theta(\tau-\mu)}. \quad . \quad (31)$$

Now let

$$\lambda_1=\sqrt{\frac{m(m+\kappa^2)}{m+1}}. \quad . \quad . \quad . \quad . \quad (32)$$

Then, using (29) and (20),

$$\psi=\left[\beta-\mathrm{D}\alpha-i\frac{2\mathrm{D}\alpha}{\lambda_1}\frac{\Theta'(\mu)}{\Theta(\mu)}\right]t+\tfrac{1}{2}\,i\log\frac{\Theta(\tau+\mu)}{\Theta(\tau-\mu)}$$

$$\equiv\omega_0 t+\tfrac{1}{2}\,i\log\frac{\Theta(\tau+\mu)}{\Theta(\tau-\mu)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

or

$$\psi=\omega_0 t+\psi_1, \quad . \quad . \quad . \quad . \quad . \quad . \quad (34)$$

* Whittaker and Watson, 'Modern Analysis,' p. 522.
† Whittaker and Watson, *loc. cit.* p. 523.

where

$$\left.\begin{aligned}\omega_0&=\beta-\mathrm{D}a-i\frac{2\mathrm{D}a}{\lambda_1}\frac{\Theta'(\mu)}{\Theta(\mu)},\\ \psi_1&=\tfrac{1}{2}\,i\log\frac{\Theta(\tau+\mu)}{\Theta(\tau-\mu)}.\end{aligned}\right\}\quad\ldots\quad(35)$$

The problem of obtaining selection rules and intensities of the asymmetrical rotator on the basis of the Bohr theory, which was the starting-point of this investigation, requires the expansion of the electric moment of the rotator, assumed to be constant, as a Fourier series in the XYZ system (fixed in space). This amounts to expanding in Fourier series the X, Y, and Z components of a constant vector (x, y, z) fixed in the moving body. It is obvious that this reduces to finding Fourier expansions of the direction cosines of the moving axes. It is more convenient, however, to use certain linear combinations of the direction cosines, resulting from the transformation

$$\left.\begin{aligned}\xi&=x+iy, & \Xi&=\mathrm{X}+i\mathrm{Y},\\ \eta&=x-iy, & \mathrm{H}&=\mathrm{X}-i\mathrm{Y},\\ \zeta&=z, & \mathrm{Z}&=\mathrm{Z}.\end{aligned}\right\}\quad\ldots\quad(36)$$

This gives the following table or matrix of coefficient of the transformation from the moving to the fixed system of axes (ξ, η, ζ are associated with the moving system, Ξ, H, Z with the fixed system).

	ξ	η	ζ
Ξ	$\frac{1+\cos\theta}{2}e^{i\phi+i\psi}$	$\frac{1-\cos\theta}{2}e^{-i\phi+i\psi}$	$-i\sin\theta\,e^{i\psi}$
H	$\frac{1+\cos\theta}{2}e^{-i\phi-i\psi}$	$\frac{1-\cos\theta}{2}e^{i\phi-i\psi}$	$i\sin\theta\,e^{-i\psi}$
Z	$-i\frac{\sin\theta}{2}e^{i\phi}$	$i\frac{\sin\theta}{2}e^{-i\phi}$	$\cos\theta$

It is to be noted that this table is to be read only horizontally, *i. e.*, it gives the quantities Ξ, H, Z in terms of ξ, η, ζ, but not the reverse.

Of the two frequencies involved in the motion, $\frac{\omega_0}{2\pi}$

is associated only with the angle ψ in the manner given in eq. (34). The expressions for $\cos\theta$, $\sin\theta$, ψ_1, and ϕ, on the other hand, involve only elliptic and Θ-functions with the same modulus, and hence have the same period, T_1. In forming expansions of the quantities given in the table, therefore, we may replace ψ by ψ_1 and expand in terms of the frequency $\frac{\omega_1}{2\pi}$ associated with the elliptic and theta functions, remembering, however, that the expansions, to be made complete, must be multiplied by $e^{i\omega_0 t}$ whenever the factor $e^{i\psi}$ is involved and by $e^{-i\omega_0 t}$ if $e^{-i\psi}$ occurs. Now from eq. (20)

$$\frac{D\alpha T_1}{\sqrt{m+1}} = 4\int_0^{\frac{\pi}{2}} \frac{d\chi}{\sqrt{1-\kappa^2\sin^2\chi}} = 4K. \quad . \quad . \quad (37)$$

K here is the complete elliptic integral of the first kind, with modulus κ *. We obtain, therefore,

$$\omega_1 = \frac{\pi D\alpha}{2K\sqrt{m+1}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

If τ is used as the time variable, the period which we may now designate τ_1 is 4K.

We now proceed to obtain Fourier series for the matrix elements given in the table by making use of eqs. (19), (21), (22), (23), and (34). We begin with

$$\frac{1+\cos\theta}{2} e^{i\phi+i\psi}.$$

We have, by eqs. (19) and (21),

$$e^{i\phi} = \cos\phi + i\sin\phi = \frac{\cos\chi}{\sqrt{1+m\sin^2\chi}} + \frac{\sqrt{m+1}\,.\,i\,.\,\sin\chi}{\sqrt{1+m\sin^2\chi}}$$
$$= \frac{\mathrm{cn}\,\tau + i\sqrt{m+1}\,\mathrm{sn}\,\tau}{\sqrt{1+m\,\mathrm{sn}^2\,\tau}}. \quad . \quad . \quad . \quad . \quad (39)$$

By making use of eqs. (22) and (34) we obtain the following expression for the matrix element as a function of τ:—

$$\tfrac{1}{2}(1+\cos\theta)e^{i\phi+i\psi_1} = \tfrac{1}{2}\left[1+\sqrt{\frac{m+\kappa^2}{m}}\,\frac{1}{\mathrm{dn}\,\tau}\right]\sqrt{\frac{\Theta(\tau-\mu)}{\Theta(\tau+\mu)}}$$
$$\times \frac{\mathrm{cn}\,\tau + i\sqrt{m+1}\,\mathrm{sn}\,\tau}{\sqrt{1+m\,\mathrm{sn}^2\,\tau}}. \quad . \quad . \quad . \quad . \quad (40)$$

* Whittaker and Watson, *loc. cit.* p. 498.

Expanding this in a Fourier series, we have

$$\tfrac{1}{2}(1+\cos\theta)\,e^{i\phi+i\psi_1}=\sum_{n=-\infty}^{\infty}\mathrm{A}_n e^{\frac{2\pi n i\tau}{4\mathrm{K}}} \quad . \quad . \quad . \quad (41)$$

where

$$\mathrm{A}_n=\frac{1}{4\mathrm{K}}\int_0^{4\mathrm{K}}\frac{1}{2}\sqrt{\frac{\Theta(\tau-\mu)}{\Theta(\tau+\mu)}}\left[1+\sqrt{\frac{m+\kappa^2}{m}}\,\frac{1}{\mathrm{dn}\,\tau}\right]$$
$$\times\frac{\mathrm{cn}\,\tau+i\sqrt{m+1}\,\mathrm{sn}\,\tau}{\sqrt{1+m\,\mathrm{sn}^2\tau}}e^{-\frac{2\pi n i\tau}{4\mathrm{K}}}\,d\tau. \quad . \quad . \quad (42)$$

Let us now make the substitution

$$\tau+2\mathrm{K}\sim\tau. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (43)$$

Since the integral is over a period, this leaves it unaltered. Further, we have *

$$\left.\begin{aligned}&\Theta(u+2\mathrm{K})=\Theta(u),\\&\mathrm{sn}\,(u+2\mathrm{K})=-\,\mathrm{sn}\,(u),\\&\mathrm{cn}\,(u+2\mathrm{K})=-\,\mathrm{cn}\,(u),\\&\mathrm{dn}\,(u+2\mathrm{K})=\mathrm{dn}\,(u).\end{aligned}\right\} \quad . \quad . \quad . \quad (44)$$

Hence we have

$$\left.\begin{aligned}&\mathrm{A}_n=(-1)^{n+1}\mathrm{A}_n\\\text{or}\\&\mathrm{A}_n=0 \text{ for } n \text{ even}.\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (45)$$

We may now expand the remaining matrix elements in the table in the same way, and show, by an argument similar to that just used, that in each case either the even or the odd terms in n vanish. We give the results below.

$$\tfrac{1}{2}(1-\cos\theta)e^{-i\phi+i\psi_1}=\sum_{n=-\infty}^{\infty}\mathrm{B}_n e^{\frac{2\pi n i}{4\mathrm{K}}\tau}, \quad . \quad . \quad (46)$$

where

$$\mathrm{B}_n=\frac{1}{4\mathrm{K}}\int_0^{4\mathrm{K}}\frac{1}{2}\sqrt{\frac{\Theta(\tau-\mu)}{\Theta(\tau+\mu)}}\left[1-\sqrt{\frac{m+\kappa^2}{m}}\,\frac{1}{\mathrm{dn}\,\tau}\right]$$
$$\times\frac{\mathrm{cn}\,\tau-i\sqrt{m+1}\,\mathrm{sn}\,\tau}{\sqrt{1+m\,\mathrm{sn}^2\tau}}e^{-\frac{2\pi n i}{4\mathrm{K}}\tau}\,d\tau. \quad . \quad . \quad (47)$$

* Whittaker and Watson, *loc. cit.* pp. 479 and 500.

Here $B_n=0$ for n even.

$$\sin\theta e^{i\psi_1}=\sum_{n=-\infty}^{\infty} C_n e^{\frac{2\pi n i}{4K}\tau}, \quad . \quad . \quad . \quad . \quad (48)$$

where

$$C_n=\frac{1}{4K}\int_0^{4K}\sqrt{\frac{\Theta(\tau-\mu)}{\Theta(\tau+\mu)}}\sqrt{\frac{1+m\,\mathrm{sn}^2\,\tau}{1-\kappa^2\,\mathrm{sn}^2\,\tau}}\sqrt{-\frac{\kappa^2}{m}}e^{-\frac{2\pi n i}{4K}\tau}\,d\tau. \quad . \quad . \quad . \quad (49)$$

Further, $C_n=0$ for n odd.

$$\frac{1+\cos\theta}{2}e^{-i\phi-i\psi_1}=\sum_{n=-\infty}^{\infty} D_n e^{\frac{2\pi n i}{4K}\tau}, \quad . \quad . \quad (50)$$

where

$$D_n=\frac{1}{4K}\int_0^{4K}\frac{1}{2}\left[1+\sqrt{\frac{m+\kappa^2}{m}}\frac{1}{\mathrm{dn}\,\tau}\right]\sqrt{\frac{\Theta(\tau+\mu)}{\Theta(\tau-\mu)}} \times\frac{\mathrm{cn}\,\tau-i\sqrt{m+1}\,\mathrm{sn}\,\tau}{\sqrt{1+m\,\mathrm{sn}^2\,\tau}}e^{-\frac{2\pi n i}{4K}\tau}d\tau, \quad . \quad . \quad (51)$$

and $D_n=0$ for n even.

$$\frac{1-\cos\theta}{2}e^{i\phi-i\psi_1}=\sum_{n=-\infty}^{\infty} E_n e^{\frac{2\pi n i}{4K}\tau}, \quad . \quad . \quad (52)$$

with

$$E_n=\frac{1}{4K}\int_0^{4K}\frac{1}{2}\left[1-\sqrt{\frac{m+\kappa^2}{m}}\frac{1}{\mathrm{dn}\,\tau}\right]\sqrt{\frac{\Theta(\tau+\mu)}{\Theta(\tau-\mu)}} \times\frac{\mathrm{cn}\,\tau+i\sqrt{m+1}\,\mathrm{sn}\,\tau}{\sqrt{1+m\,\mathrm{sn}^2\,\tau}}e^{-\frac{2\pi n i}{4K}\tau}d\tau, \quad . \quad . \quad (53)$$

and $E_n=0$ for n even.

$$\sin\theta e^{-i\psi_1}=\sum_{n=-\infty}^{\infty} F_n e^{\frac{2\pi n i}{4K}\tau}, \quad . \quad . \quad . \quad (54)$$

where

$$F_n=\frac{1}{4K}\int_0^{4K}\sqrt{-\frac{\kappa^2}{m}}\sqrt{\frac{1+m\,\mathrm{sn}^2\,\tau}{1-\kappa^2\,\mathrm{sn}^2\,\tau}}\sqrt{\frac{\Theta(\tau+\mu)}{\Theta(\tau-\mu)}}\,e^{-\frac{2\pi n i}{4K}\tau}d\tau. \quad . \quad . \quad . \quad (55)$$

Here $F_n=0$ for n odd.

$$\tfrac{1}{2}\sin\theta e^{i\phi}=\sum_{n=-\infty}^{\infty} G_n e^{\frac{2\pi n i}{4K}\tau}, \quad . \quad . \quad . \quad (56)$$

Where

$$G_n=\frac{1}{4K}\int_0^{4K}\frac{1}{2}\sqrt{-\frac{\kappa^2}{m}}\,\frac{\operatorname{cn}\tau+i\sqrt{m+1}\operatorname{sn}\tau}{\sqrt{1-\kappa^2\operatorname{sn}^2\tau}}\,e^{-\frac{2\pi n i}{4K}\tau}\,d\tau \quad . \quad . \quad . \quad (57)$$

Here $G_n=0$ for n even.

$$\tfrac{1}{2}\sin\theta e^{-i\phi}=\sum_{n=-\infty}^{\infty} H_n e^{\frac{2\pi n i}{4K}\tau}, \quad . \quad . \quad . \quad (58)$$

with

$$H_n=\frac{1}{4K}\int_0^{4K}\frac{1}{2}\sqrt{-\frac{\kappa^2}{m}}\,\frac{\operatorname{cn}\tau-i\sqrt{m+1}\operatorname{sn}\tau}{\sqrt{1-\kappa^2\operatorname{sn}^2\tau}}\,e^{-\frac{2\pi n i}{4K}\tau}\,d\tau \quad . \quad . \quad . \quad (59)$$

and $H_n=0$ for n even.

$$\cos\theta=\sum_{n=-\infty}^{\infty} I_n e^{\frac{2\pi n i}{4K}\tau}, \quad . \quad . \quad . \quad . \quad (60)$$

where

$$I_n=\frac{1}{4K}\int_0^{4K}\sqrt{\frac{m+\kappa^2}{m}}\,\frac{1}{\operatorname{dn}\tau}\,e^{-\frac{2\pi n i\tau}{4K}}\,d\tau, \quad . \quad (61)$$

and $I_n=0$ for n odd.

An inspection of the results (41) to (61) shows that the Fourier coefficients are zero for n odd for all the matrix elements involving the ζ-axis (last column in the table), while for the remaining quantities the Fourier coefficients vanish for n even. That this must be so may be visualized by a consideration of the Poinsot representation of the motion. In order to see this, the entire system of Poinsot ellipsoid and invariable plane must be rotated about the vector of total angular momentum (through the centre of gravity) with an angular velocity $-\omega_0$ as the ellipsoid rolls on the invariable plane, so that after a time $\frac{1}{2}T_1$ the point of contact between the ellipsoid and the invariable plane will occupy the same point in space that it did initially.

The importance of this for the quantum theory of the asymmetrical rotator consists in noting that if n is the quantum number associated with the frequency $\omega_1/2\pi$, we have, by the correspondence principle :

(a) If the electric moment is along the z-axis,

$$\Delta n = 0, \pm 2, \pm 4, \pm 6, \text{ etc.} \quad . \quad . \quad . \quad . \quad (62)$$

(b) If the electric moment is in the xy plane,

$$\Delta n = \pm 1, \pm 3, \pm 5, \pm 7, \text{ etc.} \quad . \quad . \quad . \quad (63)$$

It follows that, if the electric moment has a component along the z-axis and another in the xy plane, there is no restriction on Δn.

The second quantum number, specifying the total angular momentum, which may be denoted by J, is subject to the well-known rule

$$\Delta \mathrm{J} = 0, \pm 1. \quad . \quad . \quad . \quad . \quad . \quad (64)$$

This may easily be obtained from a consideration of eq. (34) and the table. These are the selection rules for the motions or quantum states of class I.

Before we state the selection rules for motions of class II., we will give a discussion of the quantum number n, which will require writing down the quantum conditions, namely :

$$\mathrm{D} = \frac{\mathrm{J}h}{2\pi}, \quad . \quad . \quad . \quad . \quad (65)$$

$$\oint \mathrm{D} \cos\theta \, d\phi = nh. \quad . \quad . \quad . \quad . \quad (66)$$

Now in motions of class II. the rôle of the x and z axes is interchanged as compared with their rôles in motions of class I. This means that in motions of class II. θ is the angle between the x and Z axes, whereas in motions of class I, it is the angle between z and Z axes. Likewise ϕ in motions of class II. is the angle between the z-axis and the line of nodes. With n as defined by eq. (66) it can be shown, as is done in the writer's previous paper *, that n suffers a discontinuity in passing

* E. E. Witmer, *loc. cit.* Note also that the J of the present paper is there designated m.

from motions of class I. to those of class II. We shall, therefore, designate the n as defined by eq. (66) for motions of class II. by a different symbol, $\bar{n}$. Strictly speaking, of course, one should not speak of a discontinuity in a sequence of integers, but it is easily seen what is meant. A more precise statement would be the following. If in eq. (66) n is regarded as capable of assuming a continuous range of values, then n suffers a discontinuity in passing from motions of one class to those of the other, J being assumed constant during this variation. As we approach class III. from the side of class I., n approaches a certain limiting value, $n^{(J)}$, for a given value of J. Likewise, if we approach class III. from the side of class II., $\bar{n}$ approaches a certain limiting value, $n^{(J)}$, for a given J. In general $n^{(J)}$ and $\bar{n}^{(J)}$ are not integral. The range of values of n and $\bar{n}$ is stated in the writer's previous paper, and it is there shown that

$$n^{(J)}+n^{(J)}=J. \quad . \quad . \quad . \quad . \quad . \quad (67)$$

It follows, therefore, that if we define

$$n=J-\bar{n} \quad . \quad . \quad . \quad . \quad . \quad . \quad (68)$$

for motions of class II., that n will take on a continuous range of values from $n=J$ for pure rotation about the z-axis to $n=0$ for pure rotation about the x-axis. The integral values in this range are the possible values of the quantum number n.

We can now state the selection rules for motions of class II. They are obtained from those of class I. by interchanging x and z in the statements and substituting $\bar{n}$ for n.

Making use of (68) and the selection rule for J, the rules for class II. become

(a) If the electric moment is along the x-axis,

for $\quad \Delta J=\pm 1$

$$\Delta n=\pm 1, \pm 3, \pm 5, \pm 7; \quad . \quad . \quad . \quad . \quad (69)$$

for $\quad \Delta J=0$

$$\Delta n=0, \pm 2, \pm 4, \pm 6. \quad . \quad . \quad . \quad . \quad (70)$$

(b) If the electric moment is in the yz plane,

for $\quad \Delta J=\pm 1$

$$\Delta n=0, \pm 2, \pm 4, \pm 6; \quad . \quad . \quad . \quad . \quad (71)$$

for $\Delta J = 0$.

$$\Delta n = \pm 1, \pm 3, \pm 5, \pm 7. \ . \ . \ . \ . \quad (72)$$

For the case of quantum transitions from states of class I. to those of class II. it seems reasonable to suppose that when the selection rules deduced from the motions of class I. agree with those deduced from the motions of class II., the selection rules in question will be valid. If, however, there is a contradiction, no selection rule can be established. For example, if the electric moment is in the direction of the z-axis, we have, from eqs. (62) and (71), the selection rule :

For $\Delta J = \pm 1$,

$\Delta n = 0, \pm 2, \pm 4, \pm 6$, etc.;

but for $\Delta J = 0$

$\Delta n =$ anything.

Since the situation in this case is complicated, we shall not attempt to state the complete selection rules.

It can be seen from this discussion that considerable difficulties confronted the Bohr theory in the case of the asymmetrical rotator. These difficulties do not appear in the new quantum theory *.

The expressions for the Fourier coefficients A_n to I_n given above can be used to obtain approximate values of the intensities of the allowed transitions by applying the correspondence principle. It does not appear desirable to do this at present, however, since Kramers and Ittmann† have obtained intensity formulas, using wave-mechanics, which have now completely superseded the original Bohr theory.

In conclusion, the writer wishes to express his thanks to the Faculty Research Committee of the University of Pennsylvania which provided funds for clerical assistance in preparing this manuscript for publication.

University of Pennsylvania,
Philadelphia, Pa., U.S.A.

* D. M. Dennison, ' Reviews of Modern Physics,' iii. p. 322 (1931).
† Kramers and Ittmann, *loc. cit.*

XXVIII. *Nuclear Synthesis and Stellar Radiation.* By HAROLD J. WALKE, *M.Sc., A.Inst.P., Department of Physics, Washington Singer Laboratories, University College, Exeter* *.

I. *Introduction.*

IT has been assumed now for several years that some form of annihilation or transmutation of matter must occur in stars to provide the heat which they continually radiate. The annihilation hypothesis, which has been upheld by Jeans †, is based on the assumption that matter is spontaneously transformed into high-frequency γ-radiation in stellar interiors, and so provides heat radiation as its wave-length is degraded by impacts with the atoms present. The process postulated to give rise to the primary γ-radiation is the mutual annihilation of an electron and a proton corresponding to the disappearance of 1·0078 mass units and the production of a quantum of high-frequency radiation of wave-length $1{\cdot}3 \times 10^{-13}$ cm. It has been pointed out by Jeans that the Klein-Nishina formula applied to Millikan's ionization-depth curve of cosmic radiation leads to the conclusion that the wave-length of the hardest component observed has just such a value, and regards this as evidence of the annihilation of protons and electrons in stars.

The most definite attempt which has been made to explain stellar radiation by means of the transmutation of matter is due to d'E. Atkinson ‡, who assumes that the source of stellar energy is in the synthesis of atomic nuclei from hydrogen. Atkinson's theory was based on the assumption that nuclei are formed from protons and negative electrons, the proton being regarded as the unit of nuclear mass.

On this view of the nature of the fundamental components of nuclei there is one obvious difficulty which any theory of atomic synthesis must overcome which has been noticed by nearly every student of the question. The simplest direct synthesis from hydrogen to any other element is the formation of helium according to the

* Communicated by Prof. F. H. Newman, D.Sc.

† Jeans, 'The Universe around Us,' p. 189.

‡ d'E. Atkinson, Astrophys. Journal, lxxiii. p. 250 (1931).

reaction $4H^+ + 2\bar{e} \rightarrow \alpha + h\nu$, and this is certainly so improbable a process (since it involves a six-body collision), and depends in any case so extremely on the density, that it cannot be regarded as playing an important part in supplying stellar energy at all.

If, however, helium can be supplied otherwise, progressive synthesis from this element onward is much less difficult to imagine. Atkinson's theory provides sources of helium by the radioactivity of Zn 60 and Be 8, and gives a qualitative description of the abundance of the elements as well as describing the principal types of stars. He assumes that protons and electrons can penetrate and remain within nuclei, synthesis up to Zn 60 occurring as the result of the continual addition of these particles.

In a star consisting of nothing but hydrogen the synthesis processes imagined could not occur. Atkinson, however, states that we are not bound to consider absolutely pure hydrogen, and says further that in any case it is probable that once in ten million years somewhere in the star a helium nucleus might result directly from a six-body collision. As soon as one nucleus is given under suitable conditions synthesis occurs and others are rapidly generated. He therefore considers that a certain "incubation period" will be necessary before enough synthesis could occur to supply the star's energy needs.

The early stages of stellar history are, then, as follows. At first the star is a diffuse mass of hydrogen of uncertain shape and temperature. As it assumes the spherical form it rises in temperature and starts radiating. The equilibrium is somewhat transient, since the energy must be supplied almost entirely by gravitational forces at first. Nevertheless, as the temperature rises Atkinson supposes that synthesis does begin. It then continues at a rate which on two grounds rises increasingly fast: first, the synthesis processes become steadily more probable owing to rising temperature and density, and, second, the supply of helium rises at an increasing rate as the number of radioactive nuclei increases.

It is the purpose of the present paper to consider these hypotheses in the light of the most recent physical experiments. It seems evident that the latest transmutation experiments show on a laboratory scale the processes leading to nuclear synthesis in stars, and confirm the view

that stellar radiation arises as a result of such synthesis. It is certain, however, that Atkinson's theory needs amendment, as will be shown below.

II. *Annihilation of Matter.*

We still know almost nothing of the conditions under which the total annihilation of matter might occur, or whether it could ever occur at all, and the hypothesis of Jeans that an electron and proton may interact in stellar interiors to be annihilated in the form of γ-radiation does not receive support from the most recent research.

Dirac *, in his "Theory of the Electron and Proton," first assumed that the holes in the distribution of negative electricity in the atom were protons, and showed that, according to his equations, the filling up of these holes or empty states of negative energy should correspond to the annihilation of an electron and a proton. Grave difficulties were introduced by the inequality of the masses of the electron and proton, and this led to ambiguity when the calculation of the probability of such a transition was considered, on account of the fact that the energy radiated by the conversion of a stationary positive electron and a stationary negative electron is $2mc^2$, whereas the energy liberated by the conversion of a stationary electron and a stationary proton should be $(m+\mathrm{M}p)c^2$. Oppenheimer † made the computation without explicit recognition of the difference in mass of electron and proton, but obtained a transition probability which was absurdly large and which was not appreciably reduced by the substitution of $(m+\mathrm{M}p)$ for $2m$ in the final formula. On the assumption that the empty negative energy states are protons this probability should have been very small, as the mutual annihilation should occur only rarely. Oppenheimer, however, obtained for the mean life of a free proton and a free electron a value of the order 10^{-10} seconds. As he states, the protons and electrons of matter are not in general free nor uniformly distributed nor at rest. But we should hardly expect their agglomeration into nuclei or atoms to reduce appreciably the mean transition probability, and it would, on this view, be impossible for matter to persist.

* Dirac, Proc. Roy. Soc. A, cxxvi. p. 360 (1930).
† Oppenheimer, Phys. Rev. xxxv. p. 929 (1930).

However, the discovery of the positive electron has removed this difficulty and has verified in a striking fashion the essential correctness of Dirac's theory. The short life of the positive electron and its disappearance by interaction with a negative electron has been conclusively demonstrated by the experiments of Thibaud *; but this process does not result in the annihilation of a proton, and the quantum produced has either 5×10^5 e.v. or $1{\cdot}0\times10^6$ e.v. energy. Thus while it has now been proved that annihilation of electron pairs does occur, this proof merely adds to the difficulties inherent in the assumption that an electron and a proton may be totally annihilated in the form of a single quantum; and these difficulties are increased when we consider the probable constitution of the proton.

It was formerly supposed that the proton was the fundamental unit of nuclear structure, it being assumed that a nucleus of mass M and atomic number Z consisted of M protons and (M–Z) negative electrons. The presence of free electrons in the nucleus has, however, always led to insuperable difficulties, and it was recently concluded by Bohr † that the theoretical treatment of electrons as such in nuclei must be abandoned. The viewpoint now generally adopted is that the electrons within nuclei must be combined in the form of stable nuclear sub-units. Until recently it was supposed that negative electrons within nuclei were bound in the form of neutrons and α-particles.

Chadwick ‡ first assumed that the neutron consisted of a close combination of a proton and an electron, the proton regarded as a simple particle being the fundamental unit of nuclear mass. But the experiments of Kurie § on the collisions between neutrons and protons and the positron radioactivity observed by Curie and Joliot ‖ and others, as well as the theoretical arguments brought forward by Chadwick ¶ in the Bakerian Lecture for 1933, suggest that the neutron is a simple particle, being the fundamental nuclear unit, the proton consisting

* Thibaud, *Ann. de la Soc. scientifique de Bruxelles*, série B, t. liv. p. 36 (1924).
† Bohr, Volta Congress (1932).
‡ Chadwick, Proc. Roy. Soc. A, cxxxvi. p. 692 (1923).
§ Kurie, Phys. Rev. xliv. p. 463 (1933).
‖ Curie and Jolit, *Comptes Rendus*, cxcviii. p. 254 (1934).
¶ Chadwick, Proc. Roy. Soc. A, cxxxvi. p. 692 (1933).

of a close combination of a neutron and positive electron. On this view the neutron may still be regarded as a combination of a proton and an electron, for as a result of interaction between these particles a neutron would be produced, the positive electron of the proton and the negative electron being annihilated in the form of γ-radiation. It is noteworthy that the γ-radiation observed by Lea * and Auger † when paraffin wax was bombarded with neutrons can be explained assuming the complex structure of the proton ‡. In addition it is significant that Fermi § has assumed in his theory of β-ray disintegration that the proton and neutron are different nuclear quantum states of the same heavy particle. While this viewpoint has much to commend it, it seems preferable, as has been pointed out by Andrade, to consider the nucleus as built up of neutrons and positive electrons, and as far as nuclear synthesis is concerned this seems essential, for it will be shown that evolution by neutron absorption and β-radioactivity is the only way one can explain the formation of the heavy elements in the stars.

It thus seems clear that "total annihilation" only occurs with pairs of electrons of opposite sign, and that when a proton and an electron interact in such a way that annihilation could occur only the electrons disappear and a neutron remains. There is at present no evidence for the annihilation of a neutron, and on this account it seems necessary to distinguish between neutrons and electrons. It seems possible, in view of the phenomena just discussed, that electrons are not material in the same sense as are neutrons; matter (that is, neutrons) and electricity (that is, electrons) are not manifestations of the same phenomenon. It will be shown in the next section that neutrons can be regarded as the units from which nuclei have been built, and that electrons are products of the kinetic energy of neutrons (or, indeed, of any moving particles in collision) or of quantum radiation of sufficiently high frequency. The mass m_0 of a neutron, though equivalent in energy to a quantum

* Lea, 'Nature,' cxxxiii. p. 24 (1934).
† Auger, *Comptes Rendus*, cxcviii. p. 365 (1934).
‡ Walke, 'Nature,' cxxxiv. p. 495 (1934).
§ Fermi, *La Ricerca Scientifica*, ii. p. 12 (1934); *Z. f. Physik*, lxxxviii. p. 161 (1934).

of frequency m_0c^2/h, will never disappear to form a quantum of γ-radiation of this frequency, for the neutron, regarded as a simple and fundamental nuclear particle, is assumed to be indestructible. Interaction between a negative electron and a proton results in the production of γ-radiation and a neutron. Both positron and electron disappear, but the neutron is not annihilated. There is at present no evidence at all of the annihilation of neutrons, and it therefore seems most satisfactory to assume a primary cosmic distribution of neutrons from which the elements have been synthesized, which has changed in configuration and distribution, but of which the total number has remained constant.

On the viewpoint put forward the creation of neutrons from radiation is as impossible a process as their annihilation in the form of quantum radiation. The experiments of Chadwick, Blackett, and Occhialini *, Anderson and Neddermeyer †, and others have shown that electron pairs are formed as a result of the interaction of γ-radiation and nuclei, exactly as suggested by the Dirac Theory of the Electron; but there is no evidence of the formation of any other material particles, and it would appear that both the creation and annihilation of material particles only takes place if such particles are pairs of electrons of opposite sign. On this view, therefore, neutrons cannot be destroyed, nor can they be created (except by liberation from protons or nuclei), and the total number of neutrons in the universe is a constant—the cosmic constant. The total number of electrons is, however, not constant, but is continually in a state of flux, creation and annihilation proceeding simultaneously. In all cases electron pairs are involved and the total number of positive and negative electrons are equal; creation of matter from radiation results in the formation of a positron and negatron; annihilation in the simultaneous disappearance of two electrons of opposite sign as quantum radiation. But the number of neutrons remains constant, and the present total energy of the universe is the same as at its creation. It is equal to the initial kinetic energy of the neutrons, though as a result of gravitational instability and the synthesis of nuclei (as will be indicated) this energy has been transformed

* Chadwick, Blackett, and Occhialini, 'Nature,' cxxxi. p. 473 (1933).
† Anderson and Neddermeyer, Phys. Rev. xliii. p. 1034 (1934).

into γ-radiation, electron pairs, and heat and light ; and the primary neutron distribution has changed its configuration entirely.

III. *Primæval Chaos.*

Jeans has shown by an exact mathematical analysis that a gravitating gas evenly distributed throughout space would, on account of gravitational instability, condense to form a series of gigantic condensations. By assuming that initially all the matter in the universe was so distributed, he was able to calculate the masses of these huge gaseous agglomerations, and showed that these were of the same order as the masses of the extra-galactic nebulæ. By applying his calculations to these condensations Jeans was able to show that they would condense further, and form stars, since the nebular spirals could be similarly gravitationally unstable, and was thus able to describe the main architecture of the universe by the gravitational instability of the original gravitating gas. It will be shown that this same process is the first cause in the synthesis of the elements, and that a description of progressive nuclear synthesis in stars may be given by considering the latest disintegration experiments involving the use of high-speed ions.

Let us consider the initial condition of the universe as imagined by Jeans, when all the substance of the present stars and nebulæ was spread uniformly throughout space—in brief, let us start from the primæval chaos from which most theories of cosmogony have started. It has formerly been assumed, since the hydrogen nucleus is much simpler than any other, that in this state the entire universe is composed of hydrogen, and, indeed, Atkinson (*loc. cit.*) assumes that the initial condition of stars is similar, and that they commence as diffuse masses of hydrogen; but it is possible that the initial state of the universe is simpler even than this. According to the viewpoint suggested in a previous paper *, and discussed in the last section, the proton is complex, consisting of a neutron and positive electron, and the creation of electrons always involves the formation of electron pairs, so that the number of positive electrons in the universe is equal to the number of negative, and thus

* Walke Phil. Mag. xvii. p. 793 (1934).

in the simplest state all nuclei will be broken down into neutrons and the electrons and positrons will be annihilated in the form of γ-radiation. The initial condition of the universe will thus consist of a uniform distribution of neutrons, the γ-radiation, due to the annihilation of the electron pairs, appearing when equilibrium is established as kinetic energy of the neutron gas, which Harkins * named "neuton." The simplest possible state of the universe on the view is an even distribution of this gas of zero atomic number, which is incapable of becoming luminous no matter to what temperature it is raised; but this gas will be gravitationally unstable, as suggested by Jeans, as the neutrons, though uncharged, are gravitating particles. Such a primary distribution of neuton will therefore condense to form huge non-luminous nebulæ, and as a result the kinetic energy of the neutrons increases and their mean free paths get smaller, collisions becoming much more frequent. As a result hydrogen is produced, this element being the first to evolve and appearing at a quite definite stage in cosmic development when the kinetic energy of the neuton atoms had reached a critical value.

To understand how this can occur it is necessary to make use of a formula due to Perrin †, who has considered the formation of electron pairs by the collision of one particle with another. He has shown that a particle of mass M_1 on impact with a stationary particle of mass M_2 will produce a pair of electrons if its kinetic energy is equal to or greater than

$$2m_0c^2\left(\frac{m_0+M_1+M_2}{M_2}\right).$$

In the case of a particle of mass M colliding with a particle of the same mass, which is much greater than that of an electron, the minimum energy is equal to

$$2m_0c^2 \,.\, 2 = 2\times 10^6 \text{ e.v.}$$

Hence when the kinetic energy of a neutron reaches 2×10^6 e.v. it may produce an electron pair on impact with another neutron at rest. When both colliding particles interact it is probable that electrons would

* Harkins, 'Nature,' p. 23 (Jan 7th, 1933).
† Perrin, *Comptes Rendus*, cxcviii. p. 1302 (1934).

be produced for smaller values of the kinetic energy. If the positron of the pair unites with one of the neutrons a proton will be formed, or if it unites with both neutrons a diplon will be produced. Thus as a result of the gravitational attraction between neutrons the extra-galactic nebulæ are formed, and in these gigantic condensations hydrogen and diplogen are produced and energy is set free corresponding to the mass defect of the proton and diplon. This energy is set free as γ-radiation, and will be absorbed in the nebula or stars which are formed in it to produce heat as its wave-length is degraded by impacts with the atoms of hydrogen and neuton now present. With the formation of hydrogen and diplogen the temperature rises, due to the nuclear energy set free, and synthesis of the heavier elements then proceeds, as will be discussed.

IV. *The Direct Synthesis of Helium.*

The difficulty formerly experienced regarding the synthesis of helium is not apparent according to the present hypothesis, as it can be produced indirectly and at comparatively low temperatures as a result of the union of two diplons. As Atkinson stated, "Granted the helium it is not difficult to suppose that many or all other known elements (and perhaps many unstable ones unknown to us) might be formed by cumulative synthesis." He supposed that initially a small number of helium nuclei were formed as a result of the coming together simultaneously of four protons and two electrons, and that then synthesis proceeded by proton and electron capture to Zn 60, which nucleus he supposed was radio-active and emitted α-particles, supplying by its radio-activity a source of helium from which synthesis on a large scale could proceed. The nucleus he considers acts as a sort of trap and cooking-pot combined, catching four protons and two electrons at each sequence, and at intervals fettering them by emitting as radiation most of the surplus mass, combining the captives into an α-particle and emitting this after a delay given by its decay constant.

It will be shown by considering the recent disintegration experiments that this hypothesis of the unknown Zn 60 isotope is unnecessary, as helium is produced continually

in the synthesis of the light elements as well as being initially synthesized at a rapid rate by two-body collisions. The atomic physicist, with his sources of high potentials and his discharge-tubes, is synthesizing elements in the same way as is occurring in stellar interiors, and the processes observed, which result in the liberation of such large amounts of energy of the order of millions of volts, indicate how the intense radiation of stars is maintained and why their temperatures are so high. Jeans has compared the stars to gigantic X-ray apparatuses. It is more probable that they resemble gigantic discharge-tubes in which synthesis and disintegration occur simultaneously, and in which stable and unstable nuclei are continuously evolved, and in which at the same time a continual supply of the nuclear building units is produced.

In the nebulæ and stars, which we have been considering, consisting almost entirely of neuton, hydrogen, and diplogen, the direct synthesis of helium will occur at comparatively low temperatures—that is, early in stellar life—as a result of the collisions between diplons. The reactions occurring will be similar to those observed by Rutherford, Oliphant and Harteck *. These investigators bombarded with diplons, ammonium chloride, and ammonium sulphate in which hydrogen had been in part replaced by diplogen. It was found that large numbers of neutrons and protons were produced, the effects being observed with a threshold voltage of 20,000 e.v. These results are interpreted by the assumption that by the union of pairs of diplons α-particles were formed, these differing from normal helium nuclei in having 23×10^6 e.v. excess energy. They are consequently unstable, and the following actions occur :—

$$D_1^2+D_1^2\rightarrow T_1^3+\overset{\nearrow}{p},$$
$$\rightarrow He_2^3+\overset{\nearrow}{n},$$

and as a result T_1^3 and He_2^3 are produced.

In the first development of the theory under consideration it was supposed (as was suggested by Dee) that in some cases the α-particles produced in a state of excess energy reverted to the normal state by emitting the

* Rutherford, Oliphant, and Harteck, Proc. Roy. Soc. A, cxliv. p. 692 (1934).

23×10^6 e.v. excess energy in the form of γ-radiation, and, indeed, it was suggested * therefore that this radiation formed a component of the primary cosmic ray beam. Recent experiments by Bleakney, Harnwell, Lozier Smith, and Smyth † indicate, however, that this action does not occur, and this is supported by the results of Oliphant and Klemperer ‡, who have failed to observe the emission of γ-radiation when heavy hydrogen is bombarded with diplons. The initial formation of helium must therefore be due to another nuclear reaction, and it is therefore suggested that He_2^4 is produced as the result of β-radioactivity of H_1^4, this unstable isotope being produced when H_1^3 captures a low-energy neutron, this missing isotope being the first synthesis link in the primary cosmic synthesis process.

It is significant that T_1^3 and He_2^3 were formed at 20,000 e.v. by bombarding diplogen at rest with diplons accelerated to this energy. This suggests that the direct collision between two diplons of energy 10,000 e.v. would produce similar results, so the formation of T_1^3 and He_2^3 could take place at stellar temperatures of the order 15,000,000° Å. In addition, as a result of such reactions very high energy protons and neutrons are produced, and these, by interacting with other atoms, cause reactions, similar to those observed in terrestrial laboratories when high-energy particles are used. It is thus apparent that on account of the above reactions nuclear transmutations may occur in stars which had previously (on energy grounds) been thought never to take place.

It is apparent too that diplogen is disintegrated by these reactions, so that, in addition to the synthesis of heavier nuclei resulting in a decrease of the amount of diplogen present, there is an increase in the amount of neuton and hydrogen, and this may be taken as typical of the cosmic evolutionary process. At least among the light elements synthesis of more complex nuclei is accompanied by an increase in the number of nuclear building units, and in consequence the possibility of equilibrium corresponding to the production

* Walke, Phil. Mag. xviii. p. 819 (1934).
† Bleakney etc., Phys. Rev. xlvi. p. 81 (1934).
‡ Oliphant, Report of the International Conference on Physics.

of protons or neutrons equal in number to those absorbed in synthesis is definite. As can be seen, however, in all the reactions so far considered the number of neutrons concerned remains constant, though they are synthesized into more complex configurations, such as protons, diplons, and α-particles.

Thus at a comparatively early stage in stellar life a mixture of neuton, hydrogen, and helium is formed, and synthesis of the heavier elements can then proceed directly.

V. *Synthesis of the Light Elements.*

R. d'E. Atkinson supposed that synthesis took place as a result of the addition of protons and electrons to nuclei. He assumed that whenever a proton penetrated a nucleus it remained inside, so that penetration and synthesis became synonymous terms. He also assumed that among the lighter elements the possibility of an electron being captured was considerable. In this connexion the most recent disintegration experiments are significant. It seems more probable that synthesis proceeds almost entirely by neutron capture, and that synthesis, by the addition of charged units such as protons, diplons, and α-particles to nuclei, produces nuclei formed in a state of excess energy, which immediately disintegrate. In fact, none of the disintegration experiments so far reported (except those involving neutrons) give evidence of the capture of the bombarding particle and the emission of excess energy in the form of γ-radiation. In every case the new nuclei produced by the addition of some charged nuclear component disintegrate and emit another component which carries away the excess nuclear energy in the form of kinetic energy of the emitted particle. The γ-radiation produced in any such reaction is emitted from the excited residual nucleus in its return to the normal state.

Experimental evidence has shown, however, that the simple capture of a neutron by a stable nucleus can occur, and we conclude, therefore, that the heavier stable isotopes of an element arise by neutron capture, the excess energy being emitted in the form of γ-rays. In certain cases of neutron addition to stable nuclei unstable nuclei have been observed, the newly formed nuclei

spontaneously transforming to an isotope of the element of atomic number greater by unity by emitting β-rays. It is suggested that these reactions explain the synthesis of nuclei by a process of neutron capture with β-radioactivity of certain missing cosmic synthesis links.

It might be objected that the experimental production of radio-nitrogen by Cockroft, Gilbert, and Walton *, who bombarded graphite with protons, is an exception to the above arguments, since it apparently indicated proton capture without the emission of any nuclear component. But the recent careful experiments of Hafstad and Tuve † have shown that the result obtained was due to diplons present in the proton stream, the excess energy in this case being associated with the emitted neutrons.

Thus the synthesis of He_2^5 (necessarily a stable isotope according to the theory postulated) could occur either by direct neutron capture or by two processes involving diplon capture, viz. :—

$$He_2^4+\overline{p+n}\rightarrow He_2^5+\overset{\nearrow}{p},$$

or

$$He_2^4+\overline{p+n}\rightarrow Li_3^5+\overset{\nearrow}{n}\quad Li_3^5\rightarrow He_2^5+\overset{+\nearrow}{e}.$$

These reactions could be tested in the laboratory by bombarding helium with high-speed diplons. Neutron capture would be indicated if proton groups were obtained, proton capture with the production of "radio-lithium" being indicated by induced positron radioactivity.

This isotope He_2^5 has not yet been detected, but it is most probably stable, consisting of an α-particle and loosely bound neutron, the latter being bound to the α-particle with almost zero binding energy. Under stellar conditions it is therefore probable that it emits the loosely bound neutron as a result either of the nuclear photoelectric effect of γ-radiation or under the bombardment by the ions present. Thus, while this isotope is probably stable, it behaves in stellar interiors as if it were radioactive, since it is so easily disintegrated into an α-particle and a neutron. The small abundance of this isotope is thus explicable, especially as He_2^4 is

* Cockroft, Gilbert, and Walton, 'Nature,' cxxxiii. p. 328 (1934).
† Hafstad and Tuve, Phys. Rev. xlv. p. 902 (1934).

continually produced by the break-up of other nuclei, so that the ratio of He_2^4 to He_2^5 is always increasing.

However, it is probable that He_2^5 can capture a low-energy neutron, the newly produced He_2^6 being spontaneously β-radioactive and transforming to Li_3^6 by the reaction

$$He_2^5 \quad \{[\alpha]+n\}+n \rightarrow He_2^6\{[\alpha]+(n+n)\}$$
$$\rightarrow Li_3^6\{[\alpha]+(\overbrace{p+n})\}+\overset{\nearrow}{\beta}.$$

This hypothesis might be tested in terrestrial laboratories by bombarding beryllium with neutrons, when the following action might occur :

$$Be_4^9 \quad \{[2\alpha]+(n)\}+n \rightarrow He_2^6\{[\alpha]+(n+n)\}+\overset{\nearrow}{\alpha},$$

and β-radioactivity of the resulting He_2^6 nucleus would be observed. Unfortunately the instability of $[2\alpha]$ is such that Be_4^9 might disintegrate into $\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}+\overset{\nearrow}{n}$ under neutron bombardment.

The synthesis of Li_3^6 probably occurs as suggested, the heavier isotope Li_3^7 being formed either by direct neutron capture or as a result of diplon absorption according to the reaction

$$Li_3^6 \quad \{[\alpha]+(\overbrace{p+n})\}+\overbrace{p+n} \rightarrow Li_3^7\{[\alpha]+(\overbrace{p+n}+n)\}+\overset{\nearrow}{p}.$$

There is therefore little difficulty in understanding the direct synthesis of the two lithium isotopes.

But the direct synthesis of beryllium introduces certain difficulties. According to Atkinson this should result from the capture of a proton by Li_3^7. Cockroft and Walton *, however, have shown that when lithium is bombarded by protons α-particles are emitted according to the action

$$Li_3^7 \quad \{[\alpha]+(\overbrace{p+n}+n)\}+p \rightarrow \overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha},$$

and, in addition, it has been found that Li_3^6 disintegrates thus :—

$$Li_3^6 \quad \{[\alpha]+(\overbrace{p+n})\}+p \rightarrow \overset{\nearrow}{\alpha}+He_2^3.$$

* Cockroft and Walton, Proc. Roy. Soc. A, cxxxvii. p. 229 (1932).

It would therefore appear that the synthesis of heavier elements by proton addition to lithium nuclei is impossible, though both actions result in the evolution of helium.

In addition, it has been shown that both lithium isotopes disintegrate after diplon capture by the following reactions :—

$$Li_3^6 \quad \{[\alpha]+(\overline{p+n})\}+\overline{p+n}\rightarrow\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha},$$

$$Li_3^6 \quad \{[\alpha]+(p+\overline{n})\}+\overline{p+n}\rightarrow Li_3^7\{[\alpha]+(\overline{p+n}+n)\}+\overset{\nearrow}{p},$$

$$Li_3^7 \quad \{[\alpha]+(p+n+n)\}+\overline{p+n}\rightarrow\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}+\overset{\nearrow}{n},$$

so that synthesis above lithium could not occur by the absorption of protons or diplons, though interaction with these particles produces a continual supply of helium. As, moreover, lithium is only disintegrated by very high-energy α-particles, resulting in the emission of neutrons, it would appear that the direct synthesis of beryllium from lithium was impossible, though boron might be formed at very high temperatures by the reaction

$$Li_3^7 \quad \{[\alpha]+(\overline{p+n}+n)\}+\alpha\rightarrow B_5^{10}\{[2\alpha]+(\overline{p+n})\}+\overset{\nearrow}{n}.$$

As, however, the neutron is a neutral particle, and can therefore penetrate all nuclei, it is probable that beryllium is directly synthesized from Li_3^7 by the absorption of a low-energy neutron, the resulting isotope Li_3^8 being radioactive and emitting β-radiation as follows :—

$$Li_3^8 \quad \{[\alpha]+(\overline{p+n}+n+n)\}\rightarrow Be_4^8\{[2\alpha]\}+\overset{\nearrow}{\beta}.$$

On Atkinson's theory Be_4^8, produced by the action $Li_3^7+p\rightarrow Be_4^8+\overset{\nearrow}{\gamma}$, is radioactive, with an average life of $10^{8\cdot5}$years, and emitting two α-particles. This assumption was made to allow the evolution of helium by a second method other than by the radioactivity of Zn 60. According to the theory adopted in this paper, however, such an assumption is unnecessary, as helium is continually evolved by the disintegration of the light elements. It is assumed therefore (as with He_2^5) that Be_4^5 is a stable, though only just stable, configuration of 2α-particles.

Its small abundance is due to the fact that at high stellar temperatures it is readily broken up into its components, since its mass defect energy is practically zero. However, the Be_4^9 isotope is formed as a result of the absorption of a neutron by this nucleus.

The most probable synthesis process for boron is then by the capture of a neutron by Be_4^9 to produce the missing radioactive isotope Be_4^{10}, which by β-ray disintegration transforms spontaneously to B_5^{10}, as follows :—

$$Be_4^{10} \quad \{[2\alpha]+(n+n)\} \rightarrow B_5^{10}\{[2\alpha]+(\overline{p+n})\}+\overrightarrow{\beta}.$$

In a similar way B_5^{11} is probably produced by the absorption of a neutron into the B_5^{10} nucleus, either by simple neutron capture or, alternatively, by the capture of a diplon by the reactions

$$B_5^{10} \quad +n \rightarrow B_5^{11},$$

$$B_6^{10} \quad \{[2\alpha]+(\overline{p+n})\}+\overline{p+n} \rightarrow B_5^{11}\{[2\alpha]+(\overline{p+n}+n)\}+\overrightarrow{p},$$

$$B_6^{10} \quad \{[2\alpha]+(\overline{p+n})\}+\overline{p+n} \rightarrow C_6^{11}\{[2\alpha]+(\overline{p+n}+p)\}+\overrightarrow{n^*},$$

$$C_6^{11} \quad \rightarrow B_5^{11}+\overrightarrow{e^+}.$$

By the second reaction protons are emitted, and in the formation of the radioactive carbon neutrons are emitted, the new nucleus in this case spontaneously emitting positive electrons.

It is noteworthy that the disintegration of lithium by protons has been observed by Von Traubenberg † at a threshold potential of 13,000 e.v. (corresponding to stellar temperatures of the order 20,000,000° Å.), or, if we assume that collisions between moving lithium atoms and protons give rise to the same reaction, at temperatures of the order 10,000,000° Å. This indicates that in stars there must exist, on account of the disintegration of lithium, very high-energy α-particles, which also play a part in the synthesis-disintegration of the elements. Consequently as the reactions :

$$D_1^2+D_1^2 \rightarrow T_1^3+\overrightarrow{p} \text{ (high energy)},$$

* Crane and Lauritsen, Phys. Rev. xlv. p. 497 (1934).
† Von Traubenberg, *Naturwiss.* xxi. p. 26 (1932).

$$D_1^2+D_1^2\rightarrow He_2^3+\overset{\nearrow}{n}\text{ (high energy)},$$

$$Li_3^7+H_1^1\rightarrow\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}\text{ (high energy)},$$

take place at comparatively low stellar temperatures all the reactions observed terrestrially by bombarding elements with high-speed protons, neutrons, and α-particles can, and most probably do, occur in stars, and the manifold of reactions which have been observed play an important part in determining the relative abundance of the elements.

The direct synthesis from boron to carbon introduces difficulties which have already been considered with lithium. By bombardment with protons and diplons both isotopes disintegrate according to the reactions

$$B_5^{11}\quad\{[2\alpha]+(\overline{p+n}+n)\}+p\rightarrow\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha},$$

$$B_5^{10}\quad\{[2\alpha]+(\overline{p+n})\}+p+n\rightarrow\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha},$$

giving further helium. It would therefore appear necessary to imagine, as with lithium, that synthesis of carbon occurs mainly by the absorption of a neutron by B_5^{11} to produce B_5^{12}, which by β-ray activity produces C_6^{12}, as follows

$$B_5^{12}\quad\{[2\alpha]+(p+n+n+n)\}\rightarrow C_6^{12}\{[3\alpha]\}+\overset{\nearrow}{\beta}.$$

Synthesis to nitrogen and C_6^{13}, however, occur by the reactions :

$$B_5^{11}\quad\{[2\alpha]+(\overline{\overline{p+n}+n})\}+\alpha\rightarrow N_7^{14}\{[3\alpha]+(\overline{p+n})\}+\overset{\nearrow}{n},$$

$$B_5^{10}\quad\{[2\alpha]+(p+\overline{n})\}+\alpha\rightarrow N_7^{13}\{[3\alpha]+p\}$$
$$\rightarrow C_6^{13}\{[3\alpha]+(n)\}+\overset{+\nearrow}{e}.$$

Carbon C_6^{12} is also produced by the capture of an α-particle by the Be_4^9 nucleus, as follows :—

$$Be_4^9\quad\{[2\alpha]+(n)\}+\alpha\rightarrow C_6^{12}\{[3\alpha]\}+\overset{\nearrow}{n}.$$

In these various ways therefore carbon is built up from the less complex nuclei, the main process being probably

as a result of the capture of a low-energy neutron by B_5^{11} the resulting nucleus being β-radioactive.

The heavier isotope C_6^{13} is probably produced by capture of a low-energy neutron, though Harkins, Gans, and Newson * have shown that by interaction with a high-speed neutron (such as is emitted in the reaction of D_1^2 on D_1^2) Be_4^9 is produced with the ejection of an α-particle. In addition C_6^{13} is produced by diplon capture according to the reactions

$$C_6^{12} \quad \{[3\alpha]\}+\overline{p+n} \rightarrow C_6^{13}\{[3\alpha]+(n)\}+\overset{\nearrow}{p},$$

$$C_6^{12} \quad \{[3\alpha]\}+\overline{p+n} \rightarrow N_6^{13}\{[3\alpha]+(p)\}+\overset{\nearrow}{n},$$

$$N_6^{13} \rightarrow C_6^{13}+\overset{+\nearrow}{e}.$$

C_6^{13} is therefore readily synthesized in stars by these reactions, though it is also formed from B_5^{10}, as already indicated.

It is formed, in addition, by the disintegration of oxygen by the reaction

$$O_8^{16} \quad \{[4\alpha]\}+n \rightarrow C_7^{13}\{[3\alpha]+(n)\}+\overset{\nearrow}{\alpha},$$

so that, as with all the lighter elements, carbon is produced both by synthesis from the lighter elements and by the disintegration of heavier ones.

Nitrogen is synthesized to a certain extent from boron by the nuclear action

$$B_5^{11} \quad \{[3\alpha]+(\overline{p+n}+n)\}+\alpha \rightarrow N_7^{14}\{[3\alpha]+(\overline{p+n})\}+\overset{\nearrow}{n},$$

though it seems most probable, since C_6^{14} is definitely missing, that the main evolutionary process for nitrogen is by the β-radioactivity of this isotope formed when C_6^{13} captures a neutron :

$$C_6^{14} \quad \{[3\alpha]+(\overline{n+n})\} \rightarrow N_7^{14}\{[3\alpha]+(\overline{p+n})\}+\overset{\nearrow}{\beta}.$$

N_7^{15} probably arises as a result of the capture of a low-energy neutron by N_7^{14}, though interaction between a high-energy neutron and N_7^{14} has been shown to produce both B_5^{11} and probably C_6^{14}, thus :—

$$N_7^{14} \quad \{[3\alpha]+(\overline{p+n})\}+n \rightarrow B_5^{11}\{[2\alpha]+(\overline{p+n}+n)\}+\overset{\nearrow}{\alpha},$$

$$N_7^{14} \quad \{[3\alpha]+(\overbrace{p+n}^{})\}+n \rightarrow C_6^{14}\{[3\alpha]+(n+n)\}+\overset{\nearrow}{p}.$$

In addition this nucleus could be synthesized by diplon capture either directly, corresponding to proton emission, or indirectly as a result of positron radioactivity of O_8^{15} by the reactions

$$N_7^{14} \quad \{[3\alpha]+(\overbrace{p+n}^{})\}+\overbrace{p+n}^{} \rightarrow N_7^{15}\{[3\alpha]+(\overbrace{p+n}^{}+n)\}+\overset{\nearrow}{p},$$

$$N_7^{14} \quad \{[3\alpha]+(\overbrace{p+n}^{})\}+\overbrace{p+n}^{} \rightarrow O_8^{15}\{[3\alpha]+(\overbrace{p+n}^{}+p)\}+\overset{\nearrow}{n},$$

$$O_8^{15} \rightarrow N_7^{15}+\overset{+\nearrow}{e}.$$

The main form of action, however, would result in the formation of C_6^{12} with the emission of α-particles. This reaction again shows how disintegration synthesis occurs with the subsequent increase in the abundance of He_2^4.

In a similar manner synthesis to oxygen proceeds mainly as a result of the reaction

$$N_7^{16} \quad \{[3\alpha]+(\overbrace{p+n}^{}+n+n)\} \rightarrow O_8^{16}\{[4\alpha]\}+\overset{\nearrow}{\beta},$$

the missing N_7^{16} (assumed the next β radioactive synthesis link) being formed by synthesis by the action :

$$N_7^{15}+n \rightarrow N_7^{16}$$

and by the disintegration of fluorine, thus :—

$$F_9^{19}+n \rightarrow N_7^{16}+\overset{\nearrow}{\alpha}.$$

The heavier isotope is then produced by simple neutron capture or by the reactions

$$N_7^{14} \quad \{[3\alpha]+(\overbrace{p+n}^{})\}+\alpha \rightarrow O_8^{17}\{[4\alpha]+(n)\}+\overset{\nearrow}{p},$$

$$O_8^{16} \quad \{[4\alpha]\}+\overbrace{p+n}^{} \rightarrow O_8^{17}\{[4\alpha]+(n)\}+\overset{\nearrow}{p},$$

$$O_8^{16} \quad \{[4\alpha]\}+\overbrace{p+n}^{} \rightarrow F_9^{17}\{[4\alpha]+p\}+\overset{\nearrow}{n},$$

$$F_9^{17} \rightarrow O_8^{17}+\overset{+\nearrow}{e}.$$

Direct interaction between O_8^{16} and a high-speed neutron (from the $D_1^2+D_1^2$ reaction) gives rise to C_6^{13}, as follows:—

$$O_8^{16} \quad \{[4\alpha]\}+n \rightarrow C_6^{13}\{[3\alpha]+(n)\}+\overset{\nearrow}{\alpha},$$

and again α-particles are emitted. Similarly O_8^{17} is itself produced by the disintegration of neon, the lighter isotope Ne_{10}^{20} emitting an α-particle on capturing the neutron.

O_8^{18} is similarly produced mainly by neutron capture, the next synthesis link being O_8^{19}, which emits β-radiation and is transformed to F_9^{19}.

It can therefore be clearly seen that the two reactions $D_1^2+D_1^2$ and $Li_3^7+H_1^1$, which can take place at accepted stellar temperatures, and which result in the emission of high-speed protons, neutrons, and α-particles, play most important parts in the synthesis of the elements. As a result of these reactions synthesis of more complex and simpler elements is proceeding at the same time. Each element is being formed by the addition of nuclear sub-units to simpler nuclei, while at the same time it is produced by the disintegration of more complex nuclei. Synthesis does not always proceed in the direction of increasing mass, as imagined by Atkinson, and, further, many radioactive isotopes are produced which emit positive or negative electrons. Protons, diplons, neutrons, and α-particles all play important parts in the gradual building up of the elements, and at the same time a continual supply of the nuclear components is produced by the reactions involved. Stellar interiors are therefore like giant discharge-tubes, the various nuclei being continually subjected to bombardment by all these particles. As a result stable and unstable nuclei are continually being formed, the stable nuclei becoming gradually more complex or simpler according to the sub-units with which they interact. The final abundance of the elements is therefore conditioned by the stability of nuclei under the bombardment, and the abundance of particular nuclei is dependent, in addition, on the stability of lighter and heavier nuclei than themselves, since, as has been shown, at least among the light elements, the building up of nuclei takes place both upward and downward in the scale of mass.

In all these actions energy is liberated which appears either as kinetic energy of the emitted particles or in the form of γ-radiation. The amount of energy set free in these actions is so great that we can readily understand why the stars are so hot and why they radiate

such enormous amounts of energy. It is to be noted, however, that although there is a continual evolution, and therefore a continual variation in the neutron configurations, the total number of neutrons remains the same. In all the reactions mass defect energy is released, but we do not need to consider neutron annihilation to provide a source of cosmic energy.

VI. *The Nature of Cosmic Radiation.*

The synthesis of the elements as discussed in this and preceding papers * is of significance in considering the nature of the primary cosmic ray-beam.

According to Millikan cosmic rays consist of electromagnetic radiations of very short wave-length. By analysis of his absorption curve he has concluded that four components of the primary beam exist, and from the results obtained has shown that the Klein-Nishina formula, combined with Aston's measurements and the Einstein formula, yield quite accurately the observed absorption coefficient of the least penetrating band if it is assumed that this component arises as a result of the synthesis of helium from hydrogen. The harder bands arise similarly by the formation of heavier elements from hydrogen.

As moreover he could detect no variation of cosmic ray intensity with sidereal time, nor any diurnal variation, and since he could detect no direct effect on the measured ionization by the Milky Way or the spiral nebula of Andromeda, he concluded that the rays must have their origin in intergalactic space and not in stars. According to him the primary cosmic radiation is photonic in nature, the quanta signalizing the formation of complex nuclei from hydrogen in the depths of space.

Jeans has suggested, however, that the most penetrating component observed by Millikan arises as a result of the complete annihilation of a proton and an electron, and has further suggested that the most recently discovered penetrating component represents the complete annihilation of an α-particle as a result of interaction with two electrons.

From the theory of nuclear evolution by neutron capture and β-radioactivity developed in this and pre-

* See note at end.

ceding papers it is now quite evident that during nuclear synthesis single units are added (by two-body encounters) to stable nuclei, and, in addition, except for neutron capture (and even in many cases with neutron capture) the excess energy is emitted not in the form of high-frequency quanta but as kinetic energy of ejected particles. In consequence, although it was previously suggested* that a quantum beam limited by an energy of 23×10^6 e.v. corresponding to the synthesis of He_2^4 by the reaction $D_1^2+D_1^2\rightarrow He_2^4+\gamma$ existed, it is now evident that during synthesis only γ-radiation of energy comparable in energy with that emitted by radioactive atoms can be produced. Thus, according to the synthesis process discussed, there can be no ultra γ-ray component of the primary cosmic radiation. Even if once during 10^{10} years two neutrons and two protons should simultaneously collide in such a way that synthesis could occur, the experimental results so far obtained indicate that T_1^3 or He_2^3 would be formed, one of the particles being emitted carrying the excess energy.

Although γ-radiation of quantum energy of the order 5×10^6 e.v. and less will be evolved in many of the proposed syntheses, it is improbable that such radiation will be emitted in any intensity from stars, for the unstable nuclei produced by neutron absorption will be formed almost entirely in stellar interiors, as Baade and Zwicky † have stated that neutrons would accumulate at the centre of a star and in consequence the evolution of the elements must proceed in stellar interiors. Thus the γ-radiation, which is the chief means whereby the nuclear energy is transformed into heat and light, has its origin well within stars, and in consequence of atomic impacts when it reaches the surface its wave-length has been increased to such an extent that it is emitted in the form of light quanta.

It is thus apparent that neither synthesis, ultra γ-radiation, nor γ-radiation of energy comparable with that of the rays emitted by radioactive atoms is emitted from stars, and such rays cannot therefore form part of the primary cosmic ray-beam.

It is, however, possible that a γ-ray component of the

* Walke, Phil. Mag. xviii. p. 795 (1934).
† Baade and Zwicky, Phys. Rev. xlvi. p. 76 (1934).

primary radiation may exist, the radiation arising, as suggested by Jeans, by the total annihilation of matter, for example, by the annihilation of a proton and an electron, the two disappearing in the form of a single quantum of radiation.

In connexion with the possibility of such a reaction the following points arise. In the first place, as was pointed out by Hughes and Jauncey *, any annihilation process must satisfy the following conditions :—

1. Consevation of energy.
2. Conservation of momentum.
3. Conservation of charge.
4. Reversibility of process considered.
5. Velocities never to exceed that of light.

It has been shown that with these conditions to be satisfied it is impossible for a photon to be formed by the mutual annihilation of a proton and an electron, no matter what their velocities or directions of motion may be. In consequence, as Eddington has observed, the annihilation of a proton and an electron (if it ever occurs) can happen only when they form part of a nucleus which will leave a residuum to carry the recoil. This could only occur where nuclei exist, *i. e.*, in the stars or nebulæ. On the other hand, as Jeans † has pointed out, the radiation could not retain its observed high penetrating power if it had already penetrated any great thickness of matter. Wherever it originated it must have passed out into intergalactic space without much absorption, which is the same as saying it must have originated in matter at comparatively low temperatures. There appears to be a fundamental discrepancy in these two conditions.

Assuming that the proton is a complex particle consisting of a neutron and positive electron, and that no negative electrons exist within the nucleus, the annihilation of a proton and an electron in a nucleus would require that a negative electron entered from outside, since if a proton entered a nucleus there would be no negative electron inside with which it could pair for annihilation. But the theoretical consideration of

* Hughes and Jauncey, ' Nature,' February 6th, 1926.
† Jeans, ' The Universe around Us.'

Furry and Carlson * and the experiments of Skobeltzyn † indicate that electrons of high energy form electron pairs on interacting with a nucleus in the same way as high-frequency γ-radiation, and the results of Skobeltzyn show that the probability of such an action is high. It therefore appears that electrons cannot enter the nucleus, so that annihilation of a proton within the nucleus is impossible, since a negative electron could not enter the nucleus to interact with it. On the viewpoint adopted interaction between an electron and a proton would result in the production of a quantum of γ-radiation due to the annihilation of the positron and incident electron, the recoil being taken up by the neutron produced.

Thus it must be concluded that no γ-ray component of the primary cosmic ray-beam exists, and it would appear that the neutron cannot be annihilated in the form of a quantum of radiation, nor can it be created by the condensation of radiation. Thus the total number of neutrons in the universe is fixed, this number being the cosmical constant.

A high-energy quantum might result from the interaction of a high-speed electron and a proton, the kinetic energy of the incident electron being shared between the recoil neutron and photon produced ‡ by the annihilation of the two electrons (though, of course, an electron of high energy on impact with a proton may produce an electron pair). If the first action is possible high-energy quanta may thus be generated in stellar atmospheres or even in the earth's atmosphere, and this appears to be the only action which could produce quanta of sufficiently high energy to form part of the cosmic ray-beam.

These conclusions are therefore such as to suggest that the incoming primary cosmic radiation is wholly corpuscular in nature, and thus cosmic rays must originate in stars and not in intergalactic space. There is, of course, considerable experimental evidence in favour of this hypothesis, for the discovery of the latitude effect alone was a strong indication of the existence of the corpuscular component (since no other completely

* Furry and Carlson, Phys. Rev. p. 236 (1933).
† Skobeltzyn, ' Nature,' cxxxiii. p. 24 (1934).
‡ See also Walke, ' Nature,' cxxxiv. p. 495 (1934).

satisfactory interpretation of the effect has been given), and with the discovery of the asymmetry the possibility of ever finding an alternative explanation practically vanishes, and we may now regard the existence of a corpuscular component as substantially proven. Moreover the studies of Johnson * show that the primary corpuscular component of the cosmic radiation at the top of the atmosphere has such an intensity that there remain at most but a few per cent. to be accounted for in some other way, such as a photon component.

VII. *Annihilation and Stellar Radiation.*

In the theory of the annihilation of atoms as the source of stellar heat, Jeans assumes that the generation of the radiation from stars proceeds spontaneously, like the disintegration of radioactive atoms. The atoms which are responsible for the light and heat of stars he suggests may be regarded as "super-radioactive" atoms which spontaneously change part or all of their substance into radiation. He therefore assumes that a star begins life with a large proportion of short-lived atoms, and so at first generates energy furiously. As it ages the shortest-lived atoms disappear first, and in so doing reduce the average energy generation, so that as a star ages so must its generation of energy decrease. Finally all the atoms with much energy-generating capacity have disappeared, and the star itself is left a shrunken and diminished mass of atoms which have very little capacity for generating radiation.

On the view that the stars contain central cores consisting largely of free neutrons these results are explicable in very similar language to that used by Jeans, since in the early life of stars there will exist a large amount of free neuton and the evolution (and disintegration) of the light elements will proceed rapidly and mainly by nuclear reactions involving neutrons. Among these elements there are many radioactive isotopes, and since the stable isotopic configurations are relatively so few in number the generation of energy is rapid, especially as the presence of large numbers of free neutrons makes the probability of the formation of radioactive nuclei very high. Thus the rapid generation of energy

* Johnson, Phys. Rev. xlv. p. 569 (1934).

by young stars is not due to " super-radioactive " atoms, but to the actual radioactive nuclei, such as H_1^4, He_2^6, Be_4^{10}, B_5^{12} etc., and to the disintegration of stable isotopes in the light element group.

As the neuton is absorbed to form stable nuclei the more complex elements are formed, and so many isotopic configurations are then possible that the probability of formation of radioactive nuclei decreases. Hence as the heavier elements are evolved and the number of free neutrons is decreased the rate of energy generation will decrease. The white dwarfs, which are stars of small mass generating very little radiation and with central temperatures of the order of hundreds of millions of degrees, are therefore probably stars in which there is only a very small number of neutrons remaining, and in which therefore the nuclear transformations by means of which mass defect energy is released ultimately as heat and light are due to the collision of α-particles and protons with nuclei. On this view therefore the white dwarfs would represent the final state of stellar evolution in which the initial distribution of neuton had assumed the maximum complexity.

Moreover on Jeans's hypothesis that matter could annihilate itself spontaneously, the rate being the same at all temperatures, there was a real difficulty in the fact that there can be no appreciable generation of energy in the substance of the earth, since the hypothesis of spontaneous annihilation at all temperatures precludes accounting for this by saying that the sun was hot and the earth cool. He therefore supposed that the interior of the sun consists of different atoms from those in the earth, and these he supposed (as already stated) to be " super-radioactive."

These facts are again in accord with the hypothesis suggested, since the concentration of neuton in stellar central regions means that there could be no free neutrons in the earth, as this was formed from the outer layers of the sun. Consequently there could be no generation of energy by the formation of radioactive isotopes by neutron capture. Moreover it is to be noted that the presence of free neutrons in stellar interiors would allow the formation of atoms of higher atomic number than uranium. Thus super-radioactive atoms may exist

in stars, and by their disintegration (but not annihilation) may therefore contribute to the energy radiated.

Finally, it is to be noted that on theory of disintegration synthesis proposed the relatively high abundance of H_1^1 compared with H_1^2 and H_1^3, and of He_2^4, relative to He_2^5, is readily explicable, since not only is H_1^2 continually being transformed to H_1^1 and H_1^3 (which, in turn, is synthesized to H_1^4, and therefore He_2^4), but H_1^1 and He_2^4 are both being continually produced by the disintegration of the light elements. Consequently the amount of H_1^1 and He_2^4 is continually increasing, and these isotopes therefore mask the other isotopes H_1^2 and H_1^3 and He_2^5, which have as yet not been detected in the stars.

Note.—This paper, in conjunction with the two preceding ones *, offer together a complete theory of nuclear synthesis by neutron capture and β-radioactivity. On the viewpoint adopted therefore the "neutron" is regarded as the fundamental nuclear component, just as the electron is the fundamental extranuclear component. On this theory therefore protons and α-particles are formed mainly within nuclei as a result of the β-radioactivity of isotopes which Aston's brilliant work has shown to be missing.

XXIX. *The Effect of Pressure on Eutectic Mixtures.*

By H. A. C. McKay *and* B. Higman †.

ALTHOUGH experimental investigations of the behaviour of two-component systems under pressure have been carried out from time to time, the theory of such systems has not yet been placed on an adequate thermodynamic basis. In particular, the equations given in the literature for the change in composition of eutectic mixtures with pressure seem to be incorrect; while the simpler pressure-temperature relation for eutectics, though obtained by several workers, seems nowhere to have been tested against the experimental data. To make such a test necessitates certain approximations. The

* Walke, Phil. Mag. xviii. p. 795 (1934); xix. p. 33 (1935).

† Communicated by the Authors.

present paper contains a derivation of the relevant equations and an attempt to test their applicability to the experimental data in view of the approximations involved.

At a binary eutectic point there are three phases in equilibrium, two solids and a liquid ; we will suppose that the solid phases are pure components and not solid solutions.

Let the two components be denoted by suffixes $_1$ and $_2$; let quantities referring to the solid phases be primed (P′, T′, etc.) and those referring to the liquid phase be unprimed. Expressing the equilibrium in terms of chemical potentials,

$$\left.\begin{aligned}\mu_1&=\mu_1'\\ \mu_2&=\mu_2'\end{aligned}\right\}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Now μ_1' and μ_2' are functions of P and T only, while μ_1 and μ_2 are functions of P, T, and N_1, where P and T are the pressure and temperature respectively, and N_1 is the mol fraction of component (1) in the liquid phase. Hence

$$\left.\begin{aligned}\frac{\partial\mu_1}{\partial T}dT+\frac{\partial\mu_1}{\partial P}dP+\frac{\partial\mu_1}{\partial N_1}dN_1&=\frac{\partial\mu_1'}{\partial T}dT+\frac{\partial\mu_1'}{\partial P}dP\\ \frac{\partial\mu_2}{\partial T}dT+\frac{\partial\mu_2}{\partial P}dP+\frac{\partial\mu_2}{\partial N_1}dN_1&=\frac{\partial\mu_2'}{\partial T}dT+\frac{\partial\mu_2'}{\partial P}dP\end{aligned}\right\}. \quad . \quad (2)$$

Now

$$\frac{\partial\mu_1}{\partial T}-\frac{\partial\mu_1'}{\partial T}=-\frac{L_1}{T},$$

where L_1 is the molar heat of solution of component (1) into the eutectic liquid. Also

$$\frac{\partial\mu_1}{\partial P}-\frac{\partial\mu_1'}{\partial P}=\overline{V}_1-V_1'=\Delta V_1,$$

where $\overline{V}_1$ is the partial molar volume of component (1) in the liquid and V_1' is the molar volume of solid component (1). Putting these, and analogous expressions for component (2), into equations (2) we obtain

$$\left.\begin{aligned}-\frac{L_1}{T}dT+\Delta V_1 dP+\frac{\partial\mu_1}{\partial N_1}dN_1&=0.\\ -\frac{L_2}{T}dT+\Delta V_2 dP+\frac{\partial\mu_2}{\partial N_1}dN_1&=0.\end{aligned}\right\}. \quad . \quad . \quad (3)$$

Thus we have two relations between the three variables used to describe the system. The system is therefore univariant.

To obtain a relation between the eutectic temperature and the pressure we eliminate the last terms of these equations by means of the Gibbs-Duhem relation in the form

$$N_1 \frac{\partial \mu_1}{\partial N_1} dN_1 + N_2 \frac{\partial \mu_2}{\partial N_1} dN_1 = 0.$$

We obtain

$$-\frac{N_1 L_1 + N_2 L_2}{T} dT + (N_1 \Delta V_1 + N_2 \Delta V_2) dP = 0,$$

$$\frac{dT}{dP} = \frac{T(N_1 \Delta V_1 + N_2 \Delta V_2)}{N_1 L_1 + N_2 L_2}. \quad . \quad . \quad . \quad (4)$$

This is the equation obtained by Lewis and Randall [1] and others.

The equation giving the change in composition of the eutectic mixture with pressure can be obtained from either of equations (3) by inserting the value of dT from equation (4). It is

$$\frac{\partial \mu_1}{\partial N_1} dN_1 = N_2 \left(\frac{L_1 \Delta V_2 - L_2 \Delta V_1}{N_1 L_1 + N_2 L_2} \right) dP. \quad . \quad . \quad . \quad (5)$$

Now

$$\mu_1 = \mu_1^0(P, T) + g_1 RT \ln N_1$$

$$\frac{\partial \mu_1}{\partial N_1} dN_1 = RT \left(\frac{\partial (g_1 \ln N_1)}{\partial N_1} \right)_{P, T} dN_1$$

$$= RT d(g_1 \ln N_1) - RT \left(\frac{\partial (g_1 \ln N_1)}{\partial P} \right)_{T, N_1} dP$$

$$- RT \left(\frac{\partial (g_1 \ln N_1)}{\partial T} \right)_{P, N_1} dT,$$

where g_1 is the osmotic coefficient of component (1) in the liquid. We shall treat the liquid phase as a semi-ideal mixture and therefore neglect the terms containing $\left(\frac{\partial g_1}{\partial P} \right)_{T, N_1}$ and $\left(\frac{\partial g_1}{\partial T} \right)_{P, N_1}$. (Semi-ideal solutions are solutions which mix at constant temperature and pressure without

absorption or emission of heat and without volume contraction or expansion [2].) Then

$$\frac{d(g_1 \ln N_1)}{dP} = \frac{N_2}{RT}\left(\frac{L_1 \Delta V_2 - L_2 \Delta V_1}{N_1 L_1 + N_2 L_2}\right). \quad . \quad . \quad (6)$$

Since the system is univariant this is a total differential equation.

This equation does not appear to have been obtained hitherto. Roloff [3], in 1895, obtained an equation of the same kind, but he omitted the osmotic coefficient. The present calculations show that it is essential to include this coefficient before applying the equation to the experimental data ; this is only to be expected, since a mixture of the eutectic composition is far from being a dilute solution.

Puschin and Grebenschtschikov [4] give, without proof, the equation

$$\frac{d \ln N_1}{dP} = -\frac{N_1 \Delta V_1 + N_2 \Delta V_2}{RT}.$$

They have not applied it to the existing data, and its validity appears doubtful, since, for example, it leads to an absurdity when we combine it with the relation

$$\frac{dN_1}{dP} + \frac{dN_2}{dP} = 0.$$

We should have

$$N_1(N_1 \Delta V_1 + N_2 \Delta V_2) + N_2(N_1 \Delta V_1 + N_2 \Delta V_2) = 0,$$

which is true only in the special case when

$$N_1 \Delta V_1 + N_2 \Delta V_2 = 0.$$

In most systems ΔV_1 and ΔV_2 are both positive.

Equation (6) can be extended in several ways. For a ternary eutectic point at which three solid phases and one liquid phase are in equilibrium, it becomes

$$\frac{d(g_1 \ln N_1)}{dP} = \frac{L_1(N_2 \Delta V_2 + N_3 \Delta V_3) - \Delta V_1(N_2 L_2 + N_3 L_3)}{RT(N_1 L_1 + N_2 L_2 + N_3 L_3)}, \quad . \quad . \quad . \quad (7)$$

and in general, for a system of n components, at the

eutectic where all n solid phases are in equilibrium with the liquid phase

$$\frac{d(g_1 \ln N_1)}{dP} = \frac{L_1 \sum_{r=1}^{n} (N_r \Delta V_r) - \Delta V_1 \sum_{r=1}^{n} (N_r L_r)}{RT \sum_{r=1}^{n} (N_r L_r)} \quad . \quad . \quad (8)$$

The equations for a binary eutectic point at which the solid phases are solid solutions instead of pure components are obtained by making the following substitutions in equations (4) and (6) :—

For L_1 write $\dfrac{x^\alpha L_1{}^\alpha + L_2{}^\alpha}{x^\alpha - x}$,

„ L_2 „ $\dfrac{x(x^\beta L_1{}^\beta + L_2{}^\beta)}{x - x^\beta}$,

„ ΔV_1 „ $\dfrac{x \Delta V_1{}^\alpha + V_{2\alpha}}{x^\alpha - x}$,

„ ΔV_2 „ $\dfrac{x(x^\beta \Delta V_1{}^\beta + \Delta V_2{}^\beta)}{x - x^\beta}$,

where α and β refer to the solid phases and the x's are mol ratios (N_1/N_2 etc.). If, finally, we do not make the approximation of treating the liquid as a semi-ideal mixture, we obtain

$$\frac{d(g_1 \ln N_1)}{dP} = \frac{L_1 - H_1{}^{dil}}{RT} \left(\frac{N_1 \Delta V_1 + N_2 \Delta V_2}{N_1 L_1 + N_2 L_2} \right) - \frac{\Delta V_1 - V_1{}^{dil}}{RT}, \quad . \quad . \quad . \quad (9)$$

where the L's and ΔV's have the same meaning as before, $H_1{}^{dil}$ is the heat of transference of one mol of component (1) from a large quantity of the eutectic liquid to pure supercooled liquid (1), and $V_1{}^{dil}$ is the volume change accompanying this process. $L_1 - H_1{}^{dil}$ is the latent heat of fusion of pure component (1) at the eutectic temperature and $\Delta V_1 - V_1{}^{dil}$ is the corresponding volume change. Unfortunately there are no experimental data available to test any of equations (7) to (9).

Eutectic points are usually discussed with reference to a freezing-point composition diagram ; this consists of two roughly linear curves which intersect at the eutectic

point. These curves may be plotted for a series of pressures, when the aggregate of eutectic points form a curve which may be called the eutectic curve in the temperature-composition diagram. (It may be worth mentioning that a pressure-composition diagram may be constructed with analogous properties.) It is with the aid of this curve that Puschin and Grebenschtschikov (*loc. cit.*) discuss their results. The equation of this curve is found by eliminating $d\mathrm{P}$ instead of $d\mathrm{T}$ from equations (3) and (4) ; it is

$$\frac{d(g_1 \ln \mathrm{N}_1)}{d\mathrm{T}} = \frac{\mathrm{N}_2}{\mathrm{RT}^2}\left(\frac{\mathrm{L}_1 \Delta \mathrm{V}_2 - \mathrm{L}_2 \Delta \mathrm{V}_1}{\mathrm{N}_1 \Delta \mathrm{V}_1 + \mathrm{N}_2 \Delta \mathrm{V}_2}\right). \quad . \quad . \quad (10)$$

This equation may also be obtained by dividing equation (6) by equation (4). Puschin and Grebenschtschikov state that this curve can never meet the axes ; this is almost certainly true, but the result does not follow from thermodynamics alone.

For the purpose of testing equation (6) we shall write it in the form

$$g_1 \mathrm{R} \frac{d\mathrm{N}_1}{d\mathrm{P}} + \mathrm{RN}_1 \ln \mathrm{N}_1 \frac{dg_1}{d\mathrm{P}} = \frac{1}{\mathrm{T}}\left(\frac{\mathrm{L}_1 \Delta \mathrm{V}_2 - \mathrm{L}_2 \Delta \mathrm{V}_1}{\frac{\mathrm{L}_1}{\mathrm{N}_2} + \frac{\mathrm{L}_2}{\mathrm{N}_1}}\right). \quad (11)$$

In this equation g_1 is obtained from the well-known expression [5] for the depression of the freezing-point of component (1),

$$g_1 \ln \frac{1}{\mathrm{N}_1} = \frac{\mathrm{L}_1{}^0}{\mathrm{RT}_1{}^0}\left(\frac{\Delta \mathrm{T}^0}{\mathrm{T}_1{}^0}\right) + \left(\frac{\mathrm{L}_1{}^0}{\mathrm{RT}_1{}^0} - \frac{\Delta \mathrm{C}_1{}^p}{2\mathrm{R}}\right)\left(\frac{\Delta \mathrm{T}^0}{\mathrm{T}_1{}^0}\right)^2. \quad (12)$$

$\mathrm{T}_1{}^0$ is the freezing-point of pure component (1) ; $\Delta \mathrm{T}^0$ is the depression of the freezing-point ; $\mathrm{L}_1{}^0$ is the latent heat of fusion of pure component (1) at its freezing-point ; and $\Delta \mathrm{C}_1{}^p = \mathrm{C}_s{}^p - \mathrm{C}_l{}^p$ is the difference between the specific heats of pure component (1) in the solid and liquid states.

The term $\mathrm{RN}_1 \ln \mathrm{N}_1 \frac{dg_1}{d\mathrm{P}}$ in equation (11) deserves special attention. The expression $\frac{dg_1}{d\mathrm{P}}$ is a total differential referring to the eutectic curve, and must not be

confused with $\left(\frac{dg_1}{dP}\right)_{T,N_1}$; for N_1 in general varies from point to point along the eutectic curve. We have already assumed $\left(\frac{dg_1}{dP}\right)_{T,N_1}$ to be zero in writing equation (6); whereas inspection of columns 7 and 8 of the table shows that the term $RN_1 \ln N_1 \frac{dg_1}{dP}$ may be more important than the term $g_1 R \frac{dN_1}{dP}$.

There are only four systems for which there exist sufficient experimental data for a test of equations (4) and (11), namely :—

(1) Diphenylamine—Urethane [4].

(2) Benzene—Urethane [6].

(3) Naphthalene—Paratoluidine [7].

(4) Naphthalene—Urethane [7].

In no other system is mention made of the composition of the eutectic under pressure. Even in these systems there are several sources of error and uncertainty :—

(*a*) The compositions are probably not known very accurately.

(*b*) The possibility of the formation of solid solutions is not discussed by any of the original investigators.

(*c*) The assumption of semi-ideality may break down. (An indication that this may occur is afforded by the naphthalene-urethane system, where a change of only 0·04 in N_1 is accompanied by a change of 0·33 in g_1.)

(*d*) The values of L, ΔV, and C^p, which are taken from the International Critical Tables and from Bridgmann [8], are known only for the pure components at their freezing-points; they are not strictly identical with the L, ΔV, and C^p in equations (4) and (11). The approximation involved here is not covered by (*c*), but involves also the neglect of the integrals $\int_{T\text{ eutectic}}^{T^0} \Delta C^p dT$ (in L)

TABLE.

System.	P (kg./cm.²).	T. °C.	N_1.	Equation (4). R.H.S.×10³.	$\frac{dT}{dP}\times 10^3$.	g_1.	$g_1 R \frac{dN_1}{dP}$.	$RN_1 \ln N_1 \frac{dg_1}{dP}$.	Equation (11). L.H.S.×10³.	R.H.S.×10³.
Diphenylamine-Urethane .	1	32·2	0·39	1·7	1·5	0·51	−5·8	1·2	−4·6	−6·9
	1550	51·6	0·175	1·0	1·1	0·45	−5·1	1·0	−4·1	−3·6
Benzene-Urethane	2550	45·9	0·51	1·1	0·8	0·41	−6·7	2·5	−4·2	−3·7
	4050	59·4	0·22	1·3	1·0	0·27	−4·5	2·4	−2·1	−2·0
Naphthalene-*p*-Toluidine .	1	30·3	0·344	3·1	2·9	0·92	0·5	0·0	0·5	−2·5
	1000	56·2	0·338	2·5	2·1	0·92	0·5	0·0	0·5	−2·1
p-Toluidine-Naphthalene..	1	30·3	0·656	..	..	0·69	−0·3	0·9	0·6	2·5
	1000	56·2	0·662	..	..	0·65	−0·3	0·9	0·6	2·1
Naphthalene-Urethane ...	1	41·2	0·221	1·6	1·2	0·50	0·8	−6·0	−5·2	−6·6
	1500	57·2	0·256	1·3	0·8	0·83	1·3	−6·2	−4·9	−5·5

In each system the component printed first is the one referred to by the suffix $_1$, and for which the osmotic data are calculated.

Owing to the absence of specific heat data no osmotic coefficients could be calculated for urethane.

and $\int_{T\ \text{eutectic}}^{T^0} \Delta\alpha\, dT$ (in ΔV), where $\Delta\alpha$ is the difference between the coefficients of expansion in the solid and liquid states.

(*e*) Terms higher than $\left(\frac{\Delta T^0}{T_1{}^0}\right)^2$ are neglected in equation (12), so there may be inaccuracies in the values of g.

(*f*) Equation (11) is particularly sensitive to small changes in N_1 and g_1.

The table shows the results of our calculations. Equation (4) is independent of (*c*), (*e*), and (*f*), so the agreement between the L.H.S. and the R.H.S. of this equation provides some sort of a check on the other approximations necessary.

In view of all the possible sources of error, the final agreement is about as good as could be expected. It is, at any rate, true to say that the experimental results are not inconsistent with our equations.

Summary.

The equations connecting the eutectic temperature, the eutectic composition, and the pressure have been deduced.

These equations have been applied to the existing experimental data and the necessary approximations discussed.

The authors wish to thank Mr. R. P. Bell for a number of valuable suggestions.

References.

(1) Lewis and Randall, ' Thermodynamics ' (New York, 1923), p. 216.
(2) Guggenheim ' Modern Thermodynamics ' (London, 1933), chap. vii.
(3) Roloff, *Z. physikal. Chem.* xvii. p. 325 (1895).
(4) Puschin and Grebenschtschikov, *Z. physikal. Chem.* cxviii. p. 276 (1925).
(5) Guggenheim, *loc. cit.* p. 110 (eq. 536).
(6) Puschin, *Z. physikal. Chem.* cxviii. p. 447 (1925). (There seems to be a misprint in Table IV. p. 452, col. 7. The curve on p. 451 makes it clear that this column should be headed " Konz. des Urethans " and not " Konz. des Benzols.")
(7) ' Tables Annuelles,' iv. p. 169 (1921), and references there.
(8) Bridgman, Physical Rev. iii. p. 126 (1924); and Proc. Amer. Acad. li. p. 55 (1916).

Balliol and Trinity College Laboratories,
Oxford.

XXX. *Dispersion and Absorption Curves for Radio Wave Propagation in the Ionosphere according to the Magneto-Ionic Theory.* *By* D. F. Martyn, *Ph.D., A.R.C.Sc.**

I. *Introduction.*

IT is now well established that the earth's magnetic field exerts considerable influence on the propagation of wireless waves in the upper atmosphere. In particular the magneto-ionic theory developed by Appleton [1], Breit [2], Goldstein [3], and Hartree [4] has achieved a number of notable successes in the interpretation of propagation phenomena. When the effects of collisions between the electrons and molecules are taken into account a frictional term is introduced into the equations of motion of the electrons, and the resulting equations for the refractive index of the medium become complex. The method of introducing this frictional term has been given by Baker and Green [5] and by Appleton and Naismith [6]. A good discussion of the resulting complex equations for the refractive indices and polarizations of the basic modes of propagation has been given by Mary Taylor [7], who has also published a number of curves illustrating the variations of these quantities with increasing ionization density N for several typical wave frequencies (p) and collision frequencies (ν), and for the particular angle made by a vertically downcoming wave with the positive direction of the earth's magnetic field at Slough, England. The derivation of such curves normally involves much labour. At the suggestion of the author this aspect has been investigated by Prof. V. A. Bailey, who has succeeded in developing an elegant graphical method [8] of making such a survey, using the device of conformal representation. The present paper gives an account of a survey which has been made by this method for three representative angles, 0°, 45°, and 90°, between the direction of propagation and the earth's field, for five typical wave-lengths (λ=100 m., 200 m., 400 m., 2000 m., 20,000 m.) and for three values of ν (10^4, 10^5, 10^6 per sec.) likely to be of practical importance in radio wave propagation. In the curves given the value of the absorption coefficient (κ) has been plotted on a

* Communicated by Prof. J. P. V. Madsen, B.Sc., B.E.

logarithmic scale down to quite small values (10^{-8}), in order to facilitate computation of the attenuation of a wave following a given path. Special attention has also been given to the limiting polarization of the basic modes when $N \to 0$, a subject on which our conclusions differ considerably from those of Dr. Taylor.

II. *Derivation of the Dispersion Curves.*

The refractive index μ, absorption coefficient κ, and polarization R are given by the following formulæ :—

$$\gamma_L^2 = \left(\alpha + i\beta - \frac{1}{c^2q^2 - 1}\right)\left(\alpha + i\beta - \frac{1}{c^2q^2 - 1} - \frac{\gamma_T^2}{1 + \alpha + i\beta}\right),$$

$$R = \frac{H_z}{H_Y} = \frac{1}{i\gamma_L}\left(\frac{1}{c^2q^2 - 1} - \alpha - i\beta\right),$$

where

$$c^2q^2 = (\mu - i\chi)^2, \qquad \chi = \kappa c/p,$$

$$\alpha = -\frac{p^2}{{p_0}^2} - \frac{1}{3},^{*} \qquad \beta = p\nu/{p_0}^2, \qquad {p_0}^2 = 4Ne^2/m,$$

$$\gamma_T = pp_T/{p_0}^2, \qquad \gamma_L = pp_L/{p_0}^2,$$

$$p_T = H_T e/mc, \qquad p_L = H_L e/mc, \qquad p_H = He/mc,$$

H_T and H_L are the components of the magnetic field perpendicular to and along the direction of propagation, c is the velocity of an electromagnetic wave *in vacuo*, $i = \sqrt{-1}$, and m, e are the mass and charge of an electron respectively.

Now Bailey has shown that if we write $R = \rho\epsilon^{-\phi}$ then ρ and ϕ may be readily determined by the aid of a chart containing a number of confocal ellipses and their associated circles. If we denote the values of the polarization constants pertaining to the "ordinary" and to the "extraordinary" basic modes by the subscripts σ and e respectively, then we have

$$\phi_\sigma = -\phi_e, \quad \text{and} \quad \rho_0 = \rho_e^{-1}.$$

The appropriate sign to be applied to ϕ_0 is determined by the quadrant in which the angle is found, while the appropriate value of ρ to be assigned to ρ_σ is determined by the fact that $\rho_0 < 1$.

* The $\frac{1}{3}$ term has been introduced by Hartree to allow for the discreteness of the scattering electrons. The necessity for its inclusion is still debatable : *v.* 'Nature,' cxxxii. p. 929-(1933).

The polarization is now completely determined, and if, following Taylor, we denote the inclination of the major axis of the ellipse of polarization to OZ by ψ, then

$$\tan 2\psi = \frac{-2\rho \cos \phi}{1-\rho^2},$$

and if the ratio of the axes of the ellipse of polarization be denoted by $\tan \theta$, then

$$\tan^2 \theta = \frac{(1+\rho^2) \sin 2\psi + 2\rho \cos \phi}{(1+\rho^2) \sin 2\psi - 2\rho \cos \phi}.$$

The sign of ψ is then determined by the sign of $(-\cos \phi)$ and that of θ by the sign of $(-\sin \phi)$. Positive values of θ correspond to right-handed polarization. Curves showing the variation of ψ_0 and θ_0 for the five frequencies corresponding to $p=2p_{\text{H}}$, p_{H}, $0{\cdot}5\,p_{\text{H}}$, $0{\cdot}1\,p_{\text{H}}$, $0{\cdot}01\,p_{\text{H}}$ (where $p_{\text{H}}=\text{He}/mc$), and for the value $\nu=10^6$/sec., are given in fig. 1. For the lower collision frequencies the curves for θ are practically unchanged, while the values of ψ become very small. It seems therefore unnecessary to reproduce the curves for these lower values of ν. The curves have been plotted for a direction of propagation making an angle of 45° with the earth's magnetic field. The polarization of the extraordinary wave may be deduced directly from the curves by virtue of the relationships

$$\theta_e = -\theta_o \text{ and } \psi_e = \psi_0 - \frac{\pi}{2}.$$

For propagation along the field the polarization is circular and $\theta_o = -45°$ for all values of N *. For propagation at right angles to the field the polarization is linear and $\psi_o = 0$ for all values of N.

Having determined ρ and ϕ it is now possible, following Bailey, to evaluate μ^2 and χ^2 with the aid of a chart containing a number of confocal parabolas. The results are shown in figs. 2*a*–2*e* for the five wave frequencies and three collision frequencies previously mentioned, and for all three angles with the magnetic field. The abscissæ are the ionization densities N, the scale being chosen in each case so as to show conveniently the complete range over which μ and χ exhibit marked variations. A fine

* This conclusion differs from that of Ratcliffe (Wireless Engineer and Exp. Wireless, x. p. 358 (1933)), who has treated the special case $\nu=0$.

Fig. 1.

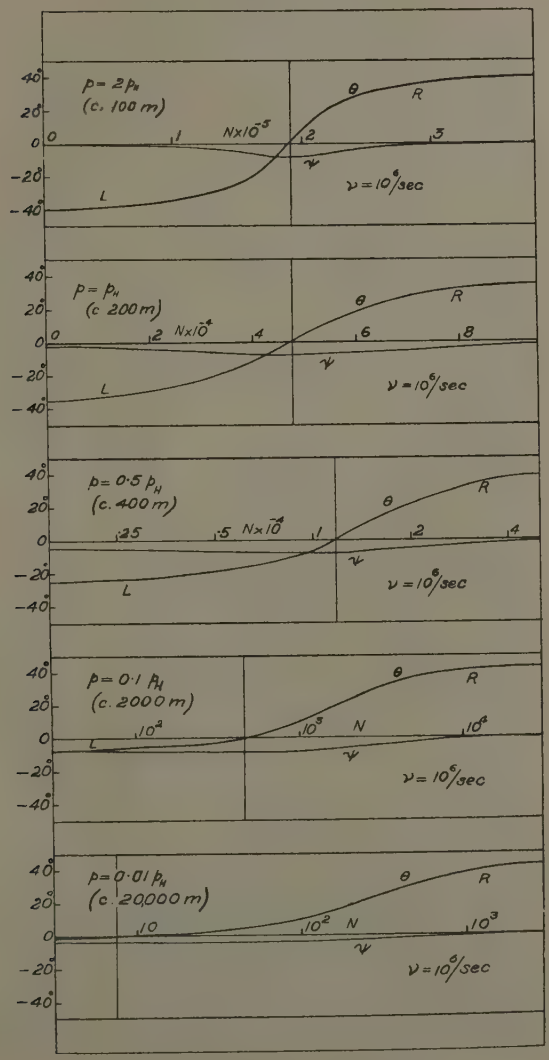

40°
20°
0
-20°
-40°
p = 2 p_H
(c. 100 m)
0
1
N x 10^-5
2
3
θ
R
ψ
L
ν = 10^6/sec
p = p_H
(c 200 m)
2
N x 10^-4
4
6
8
p = 0·5 p_H
(c. 400 m)
·25
·5
N x 10^-4
1
·1
p = 0·01 p_H
(c. 20,000 m)
10^2
10^3
N
10^4
10
10^2
10^3

Fig. 2*a*.

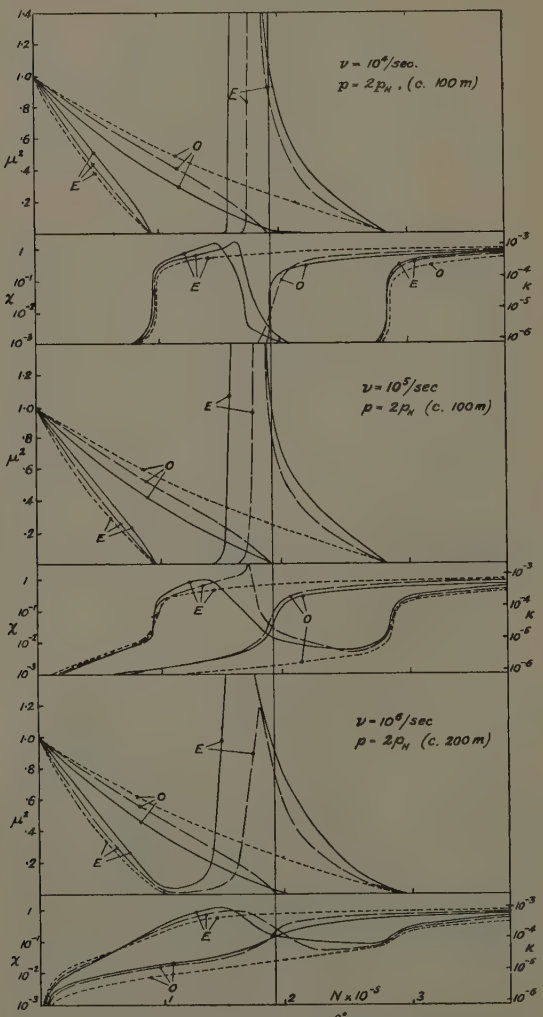

ν = 10⁴/sec.
p = 2p_H, (c. 100 m)
ν = 10⁵/sec
p = 2p_H (c. 100 m)
ν = 10⁶/sec
p = 2p_H (c. 200 m)
μ²
χ
K
O
E
N x 10⁻⁵
O – Ordinary Ray,
E – Extraordinary "
0°
45°
90°

Fig. 2 *b*.

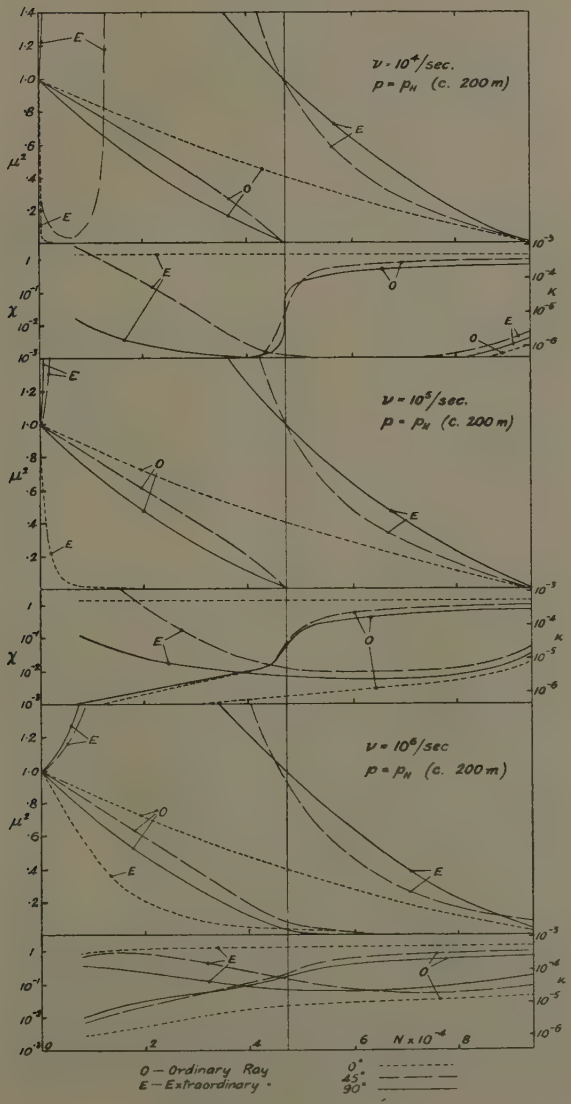

ν = 10⁴/sec.
p = pH (c. 200 m)
ν = 10⁵/sec.
p = pH (c. 200 m)
ν = 10⁶/sec
p = pH (c. 200 m)
μ²
χ
K
N x 10⁻⁴
O — Ordinary Ray
E — Extraordinary "
0°
45°
90°

Fig. 2c.

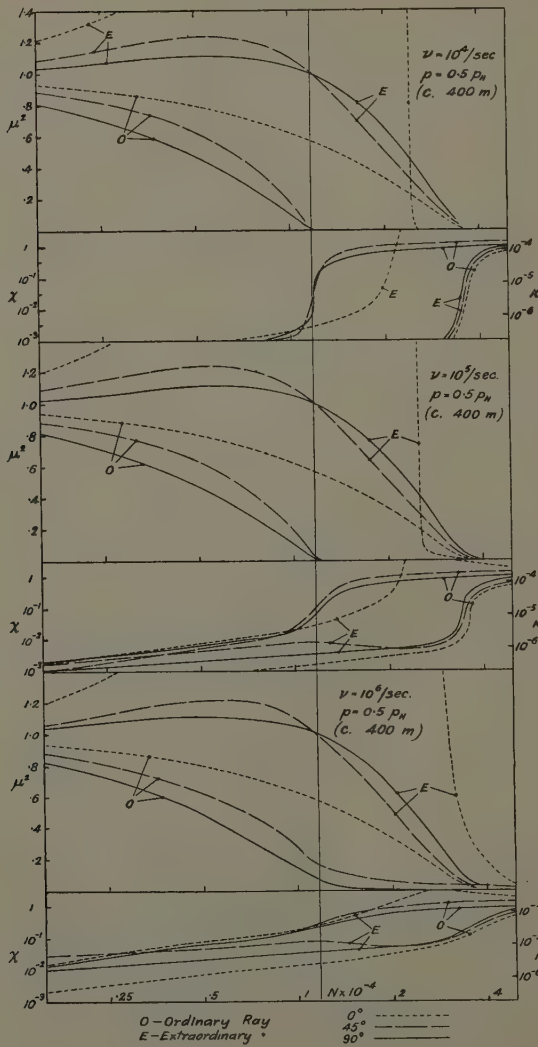

ν = 10⁴/sec
p = 0·5 pH
(c. 400 m)
ν = 10⁵/sec.
p = 0·5 pH
(c. 400 m)
ν = 10⁶/sec.
p = 0·5 pH
(c. 400 m)
μ²
χ
K
E
O
Nx10⁻⁴
O – Ordinary Ray
E – Extraordinary
0°
45°
90°

Fig. 2 *d*.

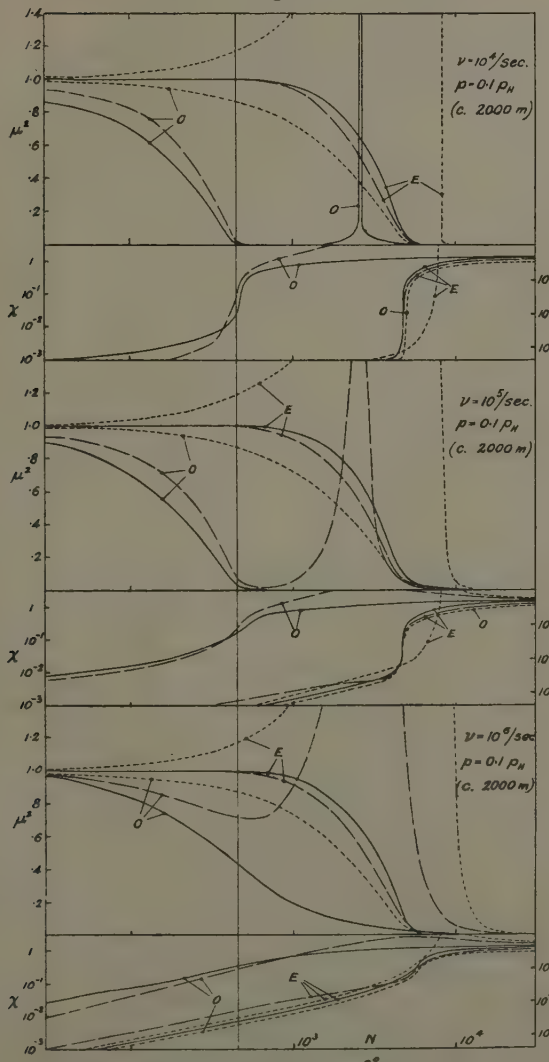
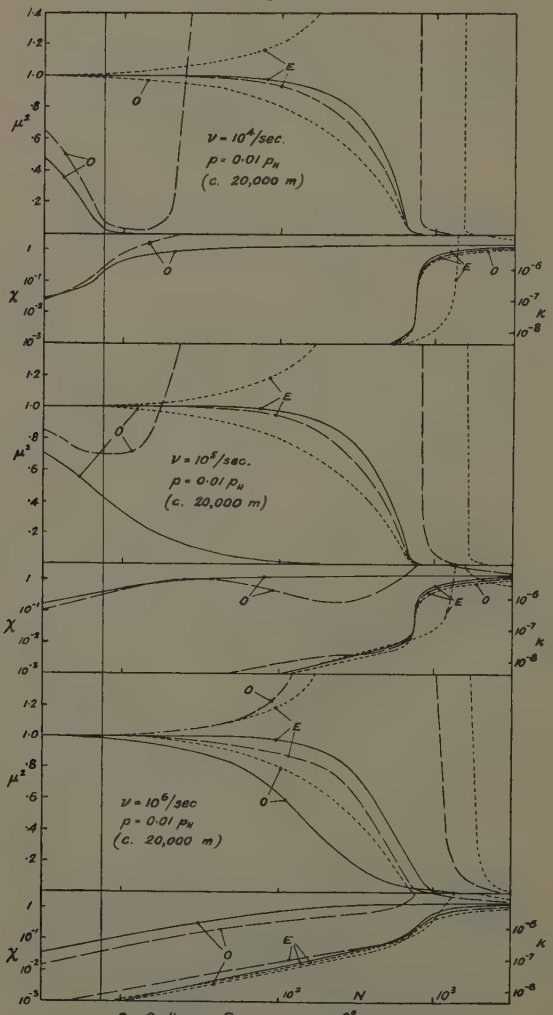

ν = 10⁴/sec.
p = 0·1 p_H
(c. 2000 m)
ν = 10⁵/sec.
p = 0·1 p_H
(c. 2000 m)
ν = 10⁶/sec.
p = 0·1 p_H
(c. 2000 m)
μ²
χ
K
N
O
E
O — Ordinary Ray
E — Extraordinary
0°
45°
60°

Fig. 2 *e*.


ν = 10⁴/sec.
p = 0·01 pH
(c. 20,000 m)
ν = 10⁵/sec.
p = 0·01 pH
(c. 20,000 m)
ν = 10⁶/sec
p = 0·01 pH
(c. 20,000 m)
μ²
χ
K
N
O
E
O – Ordinary Ray
E – Extraordinary "
0°
45°
90°

vertical line on each graph shows the value of N at which $p_0^2 = \frac{3}{2}p^2$. This is the value of ionization at which the refractive index of the ordinary ray becomes zero in the absence of collisional friction.

III. *Discussion of Curves.*

The curves of fig. 1 show that the polarization is left-handed for small values of N, linear when $p_0^2 = \frac{3}{2}p^2$, and right-handed for larger values of N, tending to circularity as $N \to \infty$. In each case the tilt of the major axis of the polarization ellipse is greatest when $p_0^2 = \frac{3}{2}p^2$, and tends to zero as $N \to \infty$. The ratio of the axes in the ellipse of polarization and the tilt of the axes both tend to limiting values as $N \to 0$. These latter results differ from those of Mary Taylor, who states (*loc. cit.*) that the polarization tends to circularity as $N \to 0$, and who gives a number of curves showing θ and ψ becoming 45° and 0° respectively for $N = 0$. Our results are, however, in accordance with those of Baker and Green [5], who show that a limiting polarization exists as the ionization decreases to zero. Moreover, Green and Builder [9] have succeeded in interpreting the long wave polarization results of Hollingworth on the basis of the existence of a limiting polarization. It would seem, therefore, that there is considerable support for our results on this point. For the longer wave-lengths the departure of the limiting polarization from circularity is pronounced, and may be so great that the linear state is approached.

It will not be necessary to discuss in detail the form of the curves given in figs. 2 *a*–2 *e*, since a general theoretical discussion has been given by Taylor [7]. It is believed that the values of collision frequency chosen for graphical illustration cover the range which is of practical importance for radio propagation in the ionosphere, *i. e.*, the range of atmospheric densities in which appreciable ionization exists. These values of ν are in general considerably lower than the critical value $\nu_c = p_T^2/2p_L$, above which Appleton and Builder [10] have shown that the type of propagation changes markedly. It follows

that the form of the dispersion curves illustrated is not greatly altered by change of ν in the region depicted except for the two longest wave-lengths.

The dispersion curves of figs. 2 *a*–2 *e* may be used in conjunction with the accompanying absorption curves to deduce information concerning the gradient of ionization in the E layer of the ionosphere and concerning the possible presence of an absorbing layer below the E layer. To accomplish this account is taken of the intensities of the downcoming waves observed in practice and of the presence or absence of the ordinary or extraordinary components in these waves. A typical example of this procedure has been given by Taylor (*loc. cit.*), who concludes that the gradient of ionization at the under surface of the E layer is sharp, that the existence of an absorbing layer below the E layer is unlikely, and that long waves are likely to be reflected at approximately the same height in the ionosphere as are the medium waves used in broadcasting. It is of interest to remark that all of these conclusions have independently been reached by the writer on quite different grounds [11]. They have again been confirmed by deduction from the curves of figs. 2 *a*–2 *e*.

IV. *Application to Long Wave Polarization Phenomena.*

Recently Green and Builder [9] have succeeded in applying the magneto-ionic theory to the interpretation of Hollingworth's measurement of the diurnal variation of the polarization of long radio waves. Owing to the complexity of the magneto-ionic equations it was necessary for these authors to make the simplifying assumption that the density of ionization was not greater than the order unity. On this assumption they deduced two rules which were of fundamental importance in their interpretation of Hollingworth's results. These rules may be put in the form :

(1) For propagation approximately transverse to the earth's magnetic field the extraordinary wave reaches the ground with much the greater intensity.

(2) If the collision frequency is great, however, approaching ν_c, then for any direction of transmission the ordinary and the extraordinary waves are equally attenuated.

Now it may be seen from consideration of fig. 2 *e* that little bending or absorption of a long radio wave occurs where N is of order unity or less. It might seem therefore that these authors' interpretation of Hollingworth's results could hardly be valid in view of the limited range of application of their rules of attenuation.

From a study of fig. 2 *e*, however, it is possible to show that the rules mentioned have a wider range of application than was proved by these authors, and in particular that they actually hold over the range of ionization densities necessary to secure the downward refraction of such waves. Thus in fig. 2 *e* it is readily seen that, when $\nu = 10^4$ or 10^5 collisions/sec., the extraordinary wave will be much less attenuated than the ordinary wave for direction of propagation making angles of 45° (or 90°) with the earth's field. On the other hand, as the collision frequency reaches the value 10^6 per sec. it is seen that the attenuation coefficients for these waves approach much nearer to equality, and, in fact, they become equal for values of ν slightly greater than 10^6 per sec. For still greater values of ν the ordinary wave may become the less attenuated, but the total attenuation is now so great that no down-coming wave is receivable, as is shown by Hollingworth's actual measurements.

It is considered therefore that the rules (1) and (2) deduced by Green and Builder are indeed valid over the complete range of N which is involved in the refraction of long waves. It seems, however, no longer desirable to infer, as do these authors, that differential absorption takes place in the post-sunrise period in a region 47 to 70 km. above the earth's surface. Hollingworth's results are now shown to be consistent with the view that the differential absorption occurs at all times in the region where the waves are bent over, that this occurs at a height of about 95 km. during the night, and at a few km. less than this height during the post-sunrise period, the fall in height being due to the descent of ionization which occurs at this period.

V. *Acknowledgments.*

This work has been carried out as part of the programme of the Radio Research Board of the Commonwealth Council for Scientific and Industrial Research, to which

the author is indebted for permission to publish his results. The author is much indebted to Prof. V. A. Bailey, of the Physics Department of the University of Sydney, for much helpful discussion; to Mr. Hutchison, of the Postmaster-General's Department, for much assistance with the plotting of the curves; to Mr. R. St. C. Pitt for assistance with the drawing of the curves; and to Prof. J. P. V. Madsen, Chairman of the Board, for his continuous encouragement and interest in the work.

VI. *Summary.*

Dispersion, absorption, and polarization curves are given for the propagation of radio waves in the ionosphere according to the magneto-ionic theory. The curves, which are obtained by Bailey's graphical method, refer to five typical wave-lengths (100 m. to 20,000 m.), three collisional frequencies (10^4, 10^5, 10^6 per sec.) likely to cover the range of practical importance for radio propagation, and for directions making the three angles 0°, 45°, and 90° with the magnetic field. It is found that the ratio of the axes of the polarization ellipses and the tilts of these axes approach limiting values as the ionization tends to zero. The curves are used to extend the conclusions of Green and Builder regarding the attenuation of long radio waves, so as to cover the full range of ionization densities penetrated by these waves. It is shown that Hollingworth's long-wave polarization measurements are consistent with the refraction and absorption of long waves at heights of about 90 km.

References.

(1) E. V. Appleton, U.R.S.I. Washington Proceedings (1927).
(2) G. Breit, Proc. Inst. Rad. Eng. xv. p. 709 (1927).
(3) S. Goldstein, Proc. Roy. Soc. A, cxxi. p. 260 (1928).
(4) D. R. Hartree, Proc. Camb. Phil. Soc. xxvii. p. 143 (1931).
(5) W. G. Baker and A. L. Green, Report no. 3, Commonwealth Radio Research Board, Bulletin 60, Australian Counc. Sci. and Indust. Res. (1932).
(6) E. V. Appleton and R. Naismith, Proc. Roy. Soc. A, cxxxvii. p. 36 (1932).
(7) M. Taylor, Proc. Phys. Soc. xlvi. p. 408 (1934).
(8) V. A. Bailey, Phil. Mag. (in press).
(9) A. L. Green and G. Builder, Proc. Roy. Soc. A, cxlv. p. 145 (1934).
(10) E. V. Appleton and G. Builder, Proc. Phys. Soc. xlv. p. 208 (1933).
(11) D. F. Martyn, "The Propagation of Medium Radio Waves in the Ionosphere," Proc. Phys. Soc. (in press).

XXXI. *On the Application of Least Squares.*—III. *A new Property of Least Squares.* *By* W. Edwards Deming, *Ph.D., Associate Physicist, Bureau of Chemistry and Soils, U.S. Department of Agriculture, Washington, D.C.**

Abstract.

It is shown that when a curve is fitted by least squares, $\Sigma(w_x U_x V_x + w_y U_y V_y) = 0$ to within higher powers of the errors, which are assumed to be small. The summation is taken over all the fitted points. w_x and V_x denote the weight and residual of the x coordinate of an observed point; U_x denotes the true error of the calculated x coordinate. w_y V_y, and U_y are similarly defined. In the simple case of observations on a single magnitude, the general proposition reduces to the familiar statement that the algebraic sum of the residuals is zero.

The distribution of χ^2 is obtained from the probability associated with ranges of errors in the observed points, the assumption being that a set of observations is a sample drawn from a normal parent population. χ^2 is defined as the sum of the squares of the x and y residuals, each residual being measured in terms of its own root mean square errror. It is pointed out that the vanishing of $\Sigma(w_x U_x V + w_y U_y V_y)$ is essential for the existence of a χ^2 distribution; hence that if the Chi-test is to be used as a criterion for testing the form of an empirical formula fitted to experimental data, it is necessary to do the fitting by least squares.

1. *Nomenclature and Definition.*

THE *error* of any *observation* is the true (unknown) value of the thing being measured subtracted from its *observed* value. Likewise the *error* of any *calculation* is the true (unknown) value of the thing being measured subtracted from its *calculated* value. For example, if

(*a*) the true value of a certain magnitude is μ;

(*b*) observations on it yield $x_1, x_2, \ldots, x_n$; then the *errors* in these observations are

$$\epsilon_1 \equiv x_1 - \mu, \quad \epsilon_2 \equiv x_2 - \mu, \quad \ldots, \quad \epsilon \equiv x_n - \mu.$$

Furthermore, if

* Communicated by the Author.

(*c*) from these observations the value of μ is estimated (calculated) to be a, then the *error* in a is

$$u \equiv a - \mu.$$

The *residual* of any observation is defined as the calculated value of the thing being measured subtracted from its observed value. Thus the *residuals* of x_1, $x_2, \ldots, x_n$ in the above example will be

$$v_1 \equiv x_1 - a, \quad v_2 \equiv x_2 - a, \ldots, \quad v_n \equiv x_n - a.$$

Errors are in practice *unknown*, but *residuals* are calculated, and, therefore, are *known*.

Table of Nomenclature.

X, Y denote the average of n_x and n_y observations on the x and y coordinates of a point.

n_x, n_y the number of observations on the x and y coordinates of a point. n_x and n_y may vary from point to point.

w_x, w_y the weights of X and Y. w_x and w_y may vary from point to point.

E_x, E_y the errors in X and Y.

x_c, y_c the calculated values of X and Y.

x_t, y_t the true values of X and Y.

U_x, U_y the errors in x_c and y_c.

V_x, V_y the residuals of X and Y.

$\alpha, \beta, \ldots$. the true values of the parameters in the empirical formula.

$a, b, \ldots$. the estimated (calculated) values of the parameters in the empirical formula.

$\mathrm{F}(x, y\,;\, a, b \ldots) = 0$ will be used for the empirical formula with its estimated parameters.

F_x^h, F_y^h, F_a^h, F_b^h denote derivatives of $\mathrm{F}(x, y\,;\, a, b, \ldots)$ evaluated at the hth point. The superscript h will be omitted when it is not desired to refer to any point in particular.

p the number of parameters in the empirical formula.

n the number of observed points.

k the number of independent residuals in the problem, or the "number of degrees of freedom." $k = n - p$.

ϕ denotes $\Sigma w v^2$ or $\Sigma(w_x \mathrm{V}_x^2 + w_y \mathrm{V}_y^2)$, the summation being taken over all observed quantities or coordinates.

σ the root mean square error of an observation of unit weight.

σ_x, σ_y the root mean square errors of single observations on the x and y coordinates of a point. σ_x and σ_y may vary from one point to another. They are related to the weights by the equations $w_x=n_x\sigma^2/\sigma_x^2$, $w_y=n_y\sigma^2/\sigma_y^2$.

χ^2 denotes ϕ/σ^2.

$\lambda_1, \lambda_2, \ldots, \lambda_n$ Lagrange multipliers introduced in previous papers in the process of minimizing ϕ.

S the slope of the empirical curve at any point. At the hth point the slope is $-F_x^h/F_y^h$.

P the probability that χ^2 will exceed a given value by the operations of chance alone.

2. *Introduction.*

It is well known that when the n observations x_1 of weight w_1, x_2 of weight w_2, $\ldots$, x_n of weight w_n have been made on a single unknown magnitude μ, the weighted mean

$$\bar{x}=\Sigma w_h x_h/\Sigma w_h$$

makes the sum of the squares of the weighted "residuals" a minimum.

Formally in this simple problem we seek to determine the parameter a in the simplest of all curves, the line

$$x=a,$$

so that

$$\phi=\Sigma w_h v_h^2=\Sigma w_h(x_h-a)^2$$

shall be a minimum. By setting $d\phi/da=0$ it is found that

$$a=\Sigma w_h x_h/\Sigma w_h=\bar{x}.$$

x is then the estimated (or calculated) value of the unknown magnitude μ, and this estimate has been made by *least squares*, for ϕ has been minimized.

The value of a just found not only minimizes ϕ, but also causes the *sum of the weighted residuals* to be zero; for a satisfies

$$\Sigma w_h(x_h-a)=0,$$

which means that

$$\Sigma w_h v_h=0.$$

The proposition contained in the equation just written is a special case of a more general one that arises in curve-fitting with least squares. The purpose of this paper

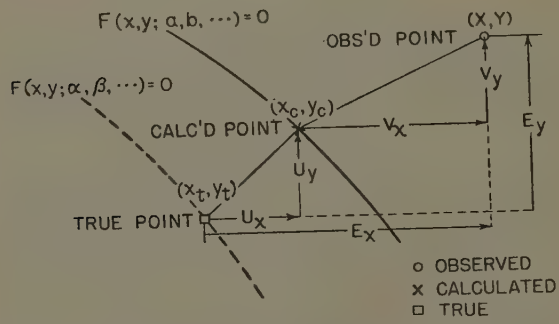
is to develop this general proposition and to indicate its importance.

3. *The probability of the Errors of the observed Points.*

The first step in developing the thesis of this paper is to examine the notion of curve-fitting and to write down the expression for the probability of a set of observed points. We may use

$$F(x, y\,;\ \alpha, \beta, \ldots)=0 \quad . \quad . \quad . \quad . \quad (66)$$

for the "true curve," which is of course in practice unknown—that is, the *true parameters* α, β, . . . are



Relations between the "true," "observed," and "calculated" points.

The x and y coordinates of a point are observed; these observations when plotted give the "observed point." The point that was measured is the "true point," which is unknown and lies on the unknown "true curve" $F(x, y\,;\ \alpha, \beta, \ldots)=0$. $\alpha, \beta, \ldots$ are the true but unknown values of the parameters. The "calculated curve" is found by adjusting a series of observed points; its equation will have the same form as the true curve, but the parameters therein will be the "calculated parameters" $a, b, \ldots$ Corresponding to each observed point there will be a "calculated point," whose coordinates are found by subtracting the "residuals" V_x and V_y from the observed coordinates X and Y. E_x and E_y denote the "errors in the observed point"; U_x and U_y denote the "errors in the calculated point." E_x, E_y, U_x, and U_y are unknown, but V_x and V_y are calculated along with the parameters a, b, ... by the method of least squares. As the figure happens to be drawn, each of the six quantities E_x, E_y, U_x, U_y, V_x, V_y is positive. Their signs are indicated by the directions of the arrows.

unknown. The *form* of the function is assumed to be known, tentatively at least; in the end the form will be retained, rejected, or placed under suspicion, depending

on how well it fits the data. A quantitative criterion for testing the goodness of fit is provided by the Chi-test, of which more will be said later.

The "calculated curve" will have the same form as the true curve, but the parameters therein will be the *calculated* or *estimated* parameters. One of the purposes of fitting a curve is to determine values for these parameters.

In the figure the x coordinate (represented by X) of the *observed point* will be the average of n_x observations on the true but unknown x coordinate (represented by x_t) of the *true point*. When the task of curve-fitting has been completed, and the observations adjusted, the *calculated* value of x (represented by x_c) can be found. Similar statements hold for y. The *true point* x_t, y_t lies on the unknown *true curve*, and the *calculated point* x_c, y_c lies on the fitted curve, the *calculated curve*. The *errors* U_x, U_y, E_x, E_y are in practice unknown, but the *residuals* V_x and V_y will be found in the process of adjustment.

Under the assumption that the normal law of error is sufficiently well obeyed *, the probability that at the point under consideration the errors in X and Y will lie respectively in the intervals $E_x \pm \frac{1}{2}dE_x$ and $E_y \pm \frac{1}{2}dE_y$ is †

$$D\,dE_x\,dE_y = \left[\frac{\sqrt{n_x}}{\sigma_x\sqrt{(2\pi)}}\,e^{-n_x E_x^2/2\sigma_x^2}\,dE_x\right] \times \left[\frac{\sqrt{n_y}}{\sigma_y\sqrt{(2\pi)}}\,e^{-n_y E_y^2/2\sigma_y^2}\,dE_y\right], \quad . \quad . \quad (67)$$

n_x and n_y being the numbers of observations on the x and y coordinates, and σ_x and σ_y the root mean square errors of single observations.

* Some remarks concerning this point will be found in an article by W. Edwards Deming and Raymond T. Birge in 'Reviews of Modern Physics,' vi. pp. 119–161 (1934).

† In this paper probability curves and surfaces will be written in differential form, and D will be used indiscriminately for the ordinates of all of them. The quantities under the differentials specify the ranges that are covered, and the whole expression represents the probability within those ranges. Thus, in eq. (67) $D\,dE_x\,dE_y$ is the probability that the errors in X and Y (the observations) lie within $E_x \pm \frac{1}{2}dE_x$ and $E_y \pm \frac{1}{2}dE_y$.

The equations in this paper will be numbered as continuations from those in two previous papers—W. Edwards Deming, Phil. Mag. xi. pp. 146–158 (1931), and xvii. pp. 804–829 (1934). These two papers will be referred to as L.S. I. and L.S. II. Equations taken from them will be given their old numbers.

So far only a single point has been taken into consideration. The situation at the other points in the problem will, however, be similar, but the errors will be entirely independent of one another. Hence, if there are n points in all, then the probability of the simultaneous occurrence of errors in the $2n$ intervals $E_{xh} \pm \frac{1}{2} dE_{xh}$, $E_{yh} \pm \frac{1}{2} dE_{yh}$ $(h=1, 2, \ldots, n)$ will be

$$D\, dE_{x1} \ldots dE_{xn}\, dE_{y1} \ldots dE_{yn} = \text{const.} \exp\left[-\tfrac{1}{2} \Sigma (n_x E_x^2/2\sigma_x^2 + n_y E_y^2/2\sigma_y^2)\right] dE_{x1} \ldots dE_{xn}\, dE_{y1} \ldots dE_{yn}. \quad . \quad (68)$$

This comes from taking the product of n factors like the one in the preceding equation. The constant in the last equation can be written out if desired; it is the product of n factors like $n_x^{\frac{1}{2}} n_y^{\frac{1}{2}} / 2\pi \sigma_x \sigma_y$—one for each point. The parenthesis following the summation is to be summed over all n points.

4. *The Transformation to Residuals.*

The exponent in the probability equation just written concerns the true but unknown errors E_x and E_y of the observed points. In order to be rendered serviceable it must be expressed in terms of residuals, since the residuals are calculated and, therefore, known. This can be accomplished by noting from the figure that at any point $E_x = V_x + U_x$, $E_y = V_y + U_y$. Hence

$$\left.\begin{aligned} E_{xh} &= V_{xh} + U_{xh} \\ E_{yh} &= V_{yh} + U_{yh} \end{aligned}\right\} h = 1, 2, \ldots, n. \quad . \quad . \quad . \quad . \quad (69)$$

The index h will be omitted when it is not necessary to refer to any point in particular.

On the left side of these transformation equations there are $2n$ independent variables. On the right there are $4n$ variables, but only $2n$ of them are independent.

If there are p independent adjustable parameters to be determined, then p pairs of errors U_x and U_y in the calculated points may be chosen at will. If we measure off the arbitrary distances $U_{x, n-p+1}$, $U_{x, n-p+2} \ldots$, U_{xn}, $U_{y, n-p+1}$, $U_{y, n-p+2}, \ldots, U_{yn}$ in the proper (i. e., x or y) direction from p arbitrarily selected "true" points on the "true curve," we shall have provided ourselves with p "calculated" points. Now these p calculated points are exactly sufficient to determine the p parameters, and hence to locate unambiguously the

"calculated curve." We may choose at will the additional $n-p$ distances U_{y1}, U_{y2}, . . . , $U_{y,n-p}$ and measure them off in the y direction from an additional $n-p$ points arbitrarily selected on the true curve; but we are *not* at liberty to choose the corresponding x distances U_{x1}, U_{x2}, . . . , $U_{x,n-p}$, for these will be determined by the calculated curve, which has already been fixed. Hence, if

$$k \equiv n-p, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (70)$$

then

$$U_{x,k+1}, U_{x,k+2}, \ldots, U_{xn}, U_{y1}, U_{y2}, \ldots, U_{yn}$$

may be considered arbitrary.

Now since the calculated curve has been fixed, and since there are p parameters, $n-p$ or k of the observed points may be placed arbitrarily. Not more than $n-p$ of the observed points may be chosen, because the observed coordinates must satisfy the normal equations, which are p in number. Ordinarily, of course, the normal equations determine the parameters, but here we are reversing the process; the parameters have already been determined, for the calculated curve was unambiguously located by the p arbitrarily chosen calculated points. This means that either the x or y residuals of k of the observed points are arbitrary. We may, therefore, choose

$$V_{x1}, V_{x2}, \ldots, V_{xk}$$

at will. The corresponding y residuals will be determinable from them, since at any point the x and y residuals are not independent of each other, but are related by

$$V_x/V_y = w_y F_x/w_x F_y = -(w_y/w_x)S = (n_y\sigma_x^2/n_x\sigma_y^2)S. \quad . \quad (71)$$

This comes from eq. (10) in L.S. I. S denotes $-F_x/F_y$, the slope of the curve at the point under consideration; F_x and F_y are written for $\partial F/\partial x$ and $\partial F/\partial y$.

It is now clear that for the $2n$ independent variables on the right we may select

$$V_{x1}, V_{x2}, \ldots, V_{xk}, U_{x,k+1}, U_{x,k+2}, \ldots, U_{xn}, U_{y1}, U_{y2}, \ldots, U_{yn}.$$

The remaining $2n$ variables on the right of eqs. (69) can be expressed in terms of these.

In transforming eq. (68) from E_x and E_y to V_x, V_y, U_x, and U_y the differential product $dE_{x1} \ldots dE_{xn}$, $dE_{y1} \ldots dE_{yn}$ is to be replaced by $J\,dV_{x1} \ldots dV_{xk}$,

$dU_{x,k+1} \ldots dU_{xn}$, $dU_{y1} \ldots dU_{yn}$, where J is the Jacobian of the $2n$ errors on the left of eqs. (69) with respect to the above $2n$ independent variables on the right. The Jacobian will be a function of the parameters, and it need not be evaluated.

To transform the exponential factor in eq. (68) we note that from eqs. (69)

$$\left.\begin{aligned} \Sigma n_x E_x^2/\sigma_x^2 &= \Sigma(n_x V_x^2/\sigma_x^2 + 2n_x U_x V_x/\sigma_x^2 + n_x U_x^2/\sigma_x^2) \\ &= (1/\sigma^2)\Sigma(w_x V_x^2 + 2w_x U_x V_x + w_x U_x^2), \\ \Sigma n_y E_y/\sigma_y^2 &= (1/\sigma^2)\Sigma(w_y V_y^2 + 2w_y U_y V_y + w_y U_y^2). \end{aligned}\right\} \quad (72)$$

It follows from these considerations that if *

$$\chi^2 \equiv \Sigma(n_x V_x^2/\sigma_x^2 + n_y V_y^2/\sigma_y^2) = (1/\sigma^2)\Sigma(w_x V_x^2 + w_y V_y^2) = \phi/\sigma^2, \quad (73)$$

where

$$\phi \equiv \Sigma(w_x V_x^2 + w_y V_y^2), \quad \ldots \quad (74)$$

then eq. (68) transformed will give

$$\begin{aligned} D\, dV_{x1} \ldots dV_{xk}\, dU_{x,k+1} \ldots dU_{yn} = \text{const. } J \exp\{-\tfrac{1}{2}[\chi^2 \\ + (2/\sigma^2)\Sigma(w_x U_x V_x + w_y U_y V_y) + (1/\sigma^2)\Sigma(w_x U_x^2 \\ + w_y U_y^2)]\}\, dV_{x1} \ldots dV_{xk}\, dU_{x,k+1} \ldots dU_{yn} \quad (75) \end{aligned}$$

for the probability that k of the x residuals lie in the intervals $V_{xh} \pm \frac{1}{2} dV_{xh}$ $(h = 1, 2, \ldots, k)$ and that the $2n-k$ independent errors of the calculated points lie in the intervals $U_{x,k+1} \pm \frac{1}{2} dU_{x,k+1}, \ldots, U_{yn} \pm \frac{1}{2} dU_{yn}$. The summations here, as elsewhere, are to be taken over all the n points.

Eq. (68) has now been transformed into an equation free of the $2n$ unknown errors E_x and E_y of the *observed* points, but inspection of eq. (75) discloses that in the process the errors U_x and U_y of the *calculated* points have been introduced, and that they make their appearance in a rather complicated manner. It is true that not all of the variables under the sign of summation are independent, but if an attempt were made to write the exponential in eq. (75) in terms of such of the variables as are independent, the problem would be even more complicated.

* χ^2 is by definition the sum of the squares of the residuals if each residual is measured in terms of its own root mean square error as a unit.

5. *The Distribution of* χ^2.

Fortunately all this complexity is of no concern for the present aims, for since the errors U_x and U_y of the calculated points are not known, the probability associated with small ranges of $U_{x,k+1}, \ldots, U_{yn}$ is of no interest. The ranges of these unknowns may therefore as well be extended from $-\infty$ to $+\infty$. This is done by integrating the right-hand side of eq. (75) with respect to each of these unknowns over the limits $-\infty$ to $+\infty$. *On the assumption that the cross term of the exponential* in eq. (75), namely,

$$\Sigma(w_x U_x V_x + w_y U_y V_y)$$

vanishes, there is left after these integrations simply

$$D\,dV_{x1} \ldots dV_{xk} = \text{const. } e^{-\frac{1}{2}\chi^2} dV_{x1} \ldots dV_{xk}, \quad . \quad (76)$$

which gives the probability that the k $(=n-p)$ independent residuals lie in the intervals $V_{xh} \pm \frac{1}{2} dV_{xh}$, $(h=1, 2, \ldots, k)$. The result of the integrations has been absorbed into the constant factor, which need not be evaluated yet.

The cross term of the exponential in eq. (75) actually *does* vanish if the curve is fitted by least squares, *i.e.*, if χ^2 is minimized. The proof of this theorem will be given in the next section.

The probability associated with the individual ranges of the k independent residuals, as expressed by eq. (76), is also of little interest. It is much more important to get the probability associated with χ^2. This, fortunately, is easy to do. By a geometrical reasoning used frequently in physics, and introduced first, so far as I know, by Karl Pearson * in 1900, it can be shown from the last equation that the distribution of χ^2 must be

$$D\,d\chi^2 = \text{const. } (\chi^2)^{\frac{1}{2}(k-2)}\, e^{-\frac{1}{2}\chi^2} d\chi^2.$$

The exponent of χ^2 is $\frac{1}{2}(k-2)$, because dimensionally the right-hand side of the last equation must be the same as the right-hand side of eq. (76), whence it was derived. The result of the integration involved in the geometrical reasoning has, like the results of the previous integrations, been absorbed into the constant multiplier.

* Karl Pearson, Phil. Mag. l. pp. 157–175 (1900). This was Pearson's development of the Chi-test for frequency distributions.

This constant multiplier in the last equation can now profitably be determined. Since χ^2 must lie between 0 and ∞, the integral of $D\,d\chi^2$ over these limits must be unity. The constant so determined and inserted back into the equation gives

$$D\,d\chi^2 = \frac{1}{\Gamma(\frac{1}{2}k)\,2^{\frac{1}{2}k}}(\chi^2)^{\frac{1}{2}(k-2)}\,e^{-\frac{1}{2}\chi^2}d\chi^2 \quad . \quad . \quad . \quad (77)$$

for the probability that χ^2 as defined by eq. (73) lies within the interval $\chi^2 \pm \frac{1}{2}\,d\chi^2$. k, the number of independent residuals in the problem, will be called the "number of degrees of freedom," to follow the terminology of R. A. Fisher.

The average value of χ^2/k is unity, for

$$\int_0^\infty D\chi^2\,d\chi^2 = k.$$

Hence, if in any problem where least squares has been used χ^2/k turns out to be so much greater than unity that one hesitates to attribute the event to chance, the indications are that the *form of the function used does not adequately describe the data*, or, what is the same thing, *constant errors are operative* at some of the points. The probability P that χ^2 will exceed a given value can be found from tables * for the Chi-test, wherein

$$P \equiv \frac{1}{\Gamma(\frac{1}{2}k)2^{\frac{1}{2}k}}\int_{\chi^2}^\infty (\chi^2)^{\frac{1}{2}(k-2)}\,e^{-\frac{1}{2}\chi^2}\,d\chi^2 \quad . \quad . \quad . \quad (78)$$

is tabulated in various ways. A very low value of P indicates that χ^2/k is significantly greater than unity. k, it will be remembered, is the number of "degrees of freedom" in the problem, being equal to $n-p$ in curve-fitting.

* The first table for using the Chi-test was published by Karl Pearson in his original memoir on the distribution of χ^2 (Phil. Mag. l. pp. 157–175 (1900)). More extensive tables were soon published by W. Palin Elderton in 'Biometrika,' i. pp. 155–163 (1901–2). Elderton's table, with additions and examples, is found in 'Tables for Statisticians and Biometricians,' part i., edited by Karl Pearson, and published by the Biometric Laboratory, University College, London, W.C. 1 (1914, 1930). Elderton's n' is our $k+1$. Another form of table is found in R. A. Fisher's 'Statistical Methods for Research Workers' (Oliver and Boyd, 1926, 4th edition, 1932). Fisher's n is the same as our k.

It is worthy of notice that the usual formulas given in textbooks for estimating σ, and hence the probable errors or root mean square errors of the observations, are equivalent to placing χ^2/k equal to unity—a practice that in the light of the foregoing developments seems extremely artificial, for it assumes that the empirical formula actually *does* fit, and that the errors are all accidental. We see now that these propositions should not be taken for granted, because they are precisely the ones that should be submitted to the Chi-test. The fact that in some problems σ may not be known, and the Chi-test consequently an impossibility, does not alter the situation.

From the integral for P it is evident that as χ^2 decreases P increases, and *vice versa*. Hence the method of least squares may be described either as the process of *minimizing* χ^2 (or ϕ) or of *maximizing* P.

When σ is known, χ^2 can be calculated from its definition in eq. (73), but in practice it should be calculated instead as a by-product of a systematic solution of the normal equations *, by which method it is unnecessary to compute the residuals V_x and V_y at any point.

The importance of calculating χ^2 and determining the goodness of fit by looking up P in any problem of curve-fitting can scarcely be overestimated. The present paper, however, is hardly the place for a lengthy discussion of the matter. The reader is referred instead to the general subject of statistical tests, which is treated in most texts on statistics, for example, R. A. Fisher's 'Statistical Methods for Research Workers' †.

The reason for introducing χ^2 here was to show that its distribution and the notion of degrees of freedom arises naturally from the probability of the occurrence of a set of errors in both x and y coordinates, when the fitting is done by least squares. The importance of least squares to the development of the χ^2 curves clearly is embodied in the vanishing of the cross term of the exponential on the right-hand side of eq. (75), the proof of which was promised for the next section.

* An example will be found in a forthcoming paper of this series.

† A remarkable series of papers on statistical tests has been presented by J. Neyman and Egon S. Pearson: 'Biometrica,' xx. A, pp. 175-241, pp. 263-294 (1928); Phil. Trans. Roy. Soc. A, ccxxxi. pp. 289-337 (1933); Proc. Camb. Phil. Soc. xxix. pp. 492-510 (1933).

6. *The Vanishing of the Cross Term—a new Property of Least Squares.*

The cross term of the exponential in eq. (75) was

$$(2/\sigma^2)\Sigma(w_x U_x V_x + w_y U_y V_x).$$

If this does not vanish, the result of integrating out the unknown errors of the calculated points—the step in passing from eq. (75) to eq. (76)—would have had a different ending; the k residuals would have appeared in eq. (76) extraneous to χ^2, and the distribution of χ^2 would have been considerably disturbed.

It is, therefore, significant that the cross term vanishes when χ^2 (or ϕ) is minimized, *i. e.*, when the method of least squares is applied. The proof can be accomplished by combining two sets of equations previously derived. In L.S. I. it was shown that

$$\Sigma\lambda_h F_a^h = 0, \quad \Sigma\lambda_h F_b^h = 0, \ldots \ldots \quad (5\,a)$$

$\lambda_1, \lambda_2, \ldots, \lambda_n$ are Lagrange multipliers that were introduced * in minimizing χ^2 (or ϕ). From eq. (10) in the same paper it is evident that

$$\lambda_h = w_{xh} V_{xh}/F_x^h = w_{yh} V_{yh}/F_y^h, \quad \ldots \ldots \quad (79)$$

a relation that will be useful shortly.

It is next to be noted that since any *true* point x_t, y_t lies on the *true* curve, and the corresponding *calculated* point x_c, y_c on the *calculated* curve,

$$F(x_t, y_t\,;\ \alpha, \beta, \ldots) = 0, \ldots \ldots \quad (80)$$

$$F(x_c, y_c\,;\ a, b, \ldots) = 0. \ldots \ldots \quad (81)$$

Now

$$x_c = x_t + U_x,\ y_c = y_t + U_y,\ a = \alpha + (a - \alpha),\ b = \beta + (b - \beta),$$

so if the errors U_x, U_y, $a - \alpha$, $b - \beta$ are all so small that their higher powers can be neglected,

$$F(x_c, y_c\,;\ a, b, \ldots) = F[x_t + U_x, y_t + U_y\,;\ \alpha + (a - \alpha),$$
$$\beta + (b - \beta), \ldots] = F(x_t, y_t\,;\ \alpha, \beta, \ldots)$$
$$+ U_x F_x + U_y F_y + (a - \alpha)F_a + (b - \beta)F_b + \ldots = 0,$$

* Actually in L.S. I. there were l additional Lagrange multipliers in eq. (5 *a*). They arose from the provision that was made for l additional relations between the parameters. But here the p parameters are supposed to be independent of one another, so the additional l Lagrange multipliers do not enter.

whence, by eq. (80),

$$U_x F_x + U_y F_y + (a-\alpha)F_a + (b-\beta)F_b + \ldots = 0. \quad . \quad (82)$$

There is a relation of this kind at every one of the n-points. We may therefore write

$$U_{xh}F_x^h + U_{yh}F_y^h + (a-\alpha)F_a^h + (b-\beta)F_b^h + \ldots = 0,$$
$$h = 1, 2, \ldots, n. \quad . \quad . \quad (83)$$

If the first of these is multiplied through by λ_1, the second by λ_2, etc., and the n equations then added together, there results, after applying eqs. (5 *a*),

$$\Sigma\lambda_h[U_{xh}F_x^h + U_{yh}F_y^h] = 0.$$

Now, by introducing the values of λ_h given in eq. (79), we have at once, dropping the index h,

$$\Sigma(w_x U_x V_x + w_y U_y V_y) = 0, \quad . \quad . \quad . \quad . \quad (84)$$

which is the vanishing of the cross term, and the thesis of this paper.

This proposition is referred to in the title as a new property of least squares. It has seemed worthy of notice because of its importance in the development and interpretation of χ^2.

The development of the distribution of χ^2 has here been carried out with the x and y coordinates both subject to error. It would be a simple matter to add a third observed coordinate, Z, at every point. To χ^2 in eq. (73) there would be appended the third term, $\Sigma n_z V_z^2/\sigma_z^2$ or $\Sigma w_z V_z^2/\sigma^2$, but the distribution of χ^2 would be left unaltered. Similarly, to the cross term of the exponential of e in eq. (75) there would be added $\Sigma w_z U_z V_z$, and it would not entail much additional work to prove that

$$\Sigma(w_x U_x V_x + w_y U_y V_y + w_z U_z V_z) = 0. \quad . \quad . \quad (85)$$

When only the x coordinate at any point is subject to error, the y and z terms for that point contribute nothing to χ^2 nor to the cross term in the exponential of eq. (75). When this is the situation at every point, χ^2 reduces to $(1/\sigma^2)\Sigma w_x V_x^2$, eq. (84) reduces to $\Sigma w_x U_x V_x = 0$, and the subscript x becomes superfluous. Similar statements hold in case only the y or the z coordinates are subject to error.

In the case of direct observations on a single unknown magnitude μ—the problem considered in the introduction—the error $u = x - \mu$ does not vary from one observation

to another ; eq. (84) then reduces to the well-known fact that $\Sigma w_h v_h = 0$, as was found in the introduction.

It seems to be well recognized that when a straight line has been fitted to observed points of which the y coordinates are of equal weight, and alone subject to error, the algebraic sum of the vertical segments connecting the observed points with the calculated line (in other words, the algebraic sum of the y residuals) is zero. More generally for the straight line, $\Sigma w_x V_x = \Sigma w_y V_y = 0$. These statements are applications of eqs. (5 *a*), but hardly of the general proposition contained in eq. (84).

XXXII. *The Sum of a Series of Cosecants.* *By* W. Ll. G. Williams.*

1. In a paper with the above title (Phil. Mag. ser. 6, xxxi. pp. 111–118 (1916)) Professor G. N. Watson has developed an asymptotic formula for the sum

$$S_n = \frac{1}{\sin\frac{\pi}{n}} + \frac{1}{\sin\frac{2\pi}{n}} + \ldots\ldots + \frac{1}{\sin(n-1)\frac{\pi}{n}}.$$

The object of the present note is to point out that this formula is readily obtained by the application of the Euler-Maclaurin sum formula to a properly chosen function. The function which we have used is one that Pólya and Szegö have used in their most interesting collection of problems † to obtain a less accurate solution of the same problem directly from the theory of Riemann Integration.

2. Following Blissard and Lucas, we use the notation B_i ($i=0, 1, 2, \ldots$) for the Bernoulli numbers, defined by

$$\frac{x}{e^x - 1} = B_0 + B_1 x + B_2 \frac{x^2}{2!} + \ldots\ldots,$$

* Communicated by the Author.

† Pólya und Szegö, ' Aufgaben und Lehrsätze aus der Analysis,' vol. i. p. 196.

or by the equivalent symbolic recursion formula *

$$(B+1)^n - B^n = 0, \ (n=2, 3, 4, \ldots),$$

where the symbol $B^n = B_n$. Thus, when $n=2$,

$$(B+1)^2 - B^2 = 2B+1 = 2B_1+1 = 0,$$

whence $B_1 = -\frac{1}{2}$. The first six of the Bernoulli numbers are

$$B_0=1, \ B_1=-1/2, \ B_2=1/6, \ B_3=0, \ B_4=-1/30,$$
$$B_5=0, \ B_6=1/42;$$

the B's with uneven subscripts, with the exception of B_1, are all zero.

3. We write the Euler-Maclaurin sum formula in the form

$$\sum_{j=0}^{n-1} f(a+jh) = \frac{1}{h}\int_a^b f(x)\,dx + \sum_{j=1}^{2i-1} \frac{B_j}{j!} h^{j-1} \triangle f^{j-1}(x) + R,$$

where $f^0(x)=f(x), f^1(x)=f'(x), f^2(x)=f''(x), \ldots;$

$$h = \frac{b-a}{n}, \ R = \frac{n\mu B_{2i} h^{2i}}{(2i)!}, \ \triangle f^{j-1}(x) = f^{j-1}(b) - f^{j-1}(a),$$

μ being a mean value of $f^{2i}(x)$. And we apply this in the interval $0 \leq x \leq \pi$ to the function

$$\begin{cases} f(x) = \dfrac{1}{\sin x} - \dfrac{1}{x} - \dfrac{1}{\pi - x}, & 0 < x < \pi, \\ f(0) = f(\pi) = \lim\limits_{x \to 0} \left(\dfrac{1}{\sin x} - \dfrac{1}{x} - \dfrac{1}{\pi - x} \right) = \dfrac{-1}{\pi}. \end{cases}$$

In view of the known expansion, valid for $0 < |x| < \pi$,

$$\frac{x}{\sin x} = 1 + \sum_{j=1}^{\infty} (-1)^{j+1} \frac{B_{2j}}{(2j)!} (2^{2j}-2)\, x^{2j},$$

we have for $0 \leq x < \pi$

$$f(x) = -\frac{1}{\pi - x} + \sum_{j=1}^{\infty} (-1)^{j+1} \frac{B_{2j}}{(2j)!} (2^{2j}-2) x^{2j-1}.$$

This formula and the relation $f(x)=f(\pi-x)$ show that $f(x)$ possesses derivatives of all orders in the closed interval

* *Cf.* de la Vallée-Poussin, 'Cours d'Analyse' (fourth edition), vol. ii., or Nörlund's "Mémoire sur les Polynomes de Bernoulli," Acta Math. vol. xliii (1922).

$0 \leq x \leq \pi$, and hence that each of these derivatives is bounded in the interval under consideration. Since all the Bernoulli numbers with uneven subscripts, except B_1, are zero, we require only

$$\begin{aligned}\triangle f^{2j-1}(x) &= f^{2j-1}(\pi) - f^{2j-1}(0) = -2f^{2j-1}(0) \\ &= 2\left\{ \frac{(2j-1)!}{\pi^{2j}} + (-1)^j \frac{B_{2j}}{2j}(2^{2j}-2) \right\}.\end{aligned}$$

Thus the Euler-Maclaurin sum formula gives us, after elementary calculations which we suppress,

$$\pi S_n = 1 + 2n(\log 2 - \log \pi) + \sum_{j=1}^{i-1} (-1)^j \frac{B^2_{2j}}{j(2j)!} \frac{\pi^{2j}}{n^2}(2^{2j}-2)$$

$$+ 2n\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1} + \frac{B_2}{2n^2} + \ldots + \frac{B_{2i-2}}{(2i-2)\, n^{2i-2}} \right.$$

$$\left. + \frac{\theta B_{2i}}{2i\, n^{2i}} \right) + R - \frac{2n\theta B_{2i}}{2i\, n^{2i}}, \text{ where } R = \frac{\mu B_{2i} \pi^{2i}}{(2i)!\, n^{2i-1}}$$

and θ is so chosen as a number between 0 and 1 that *

$$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1} + \frac{B_2}{2n^2} + \ldots + \frac{B_{2i-2}}{(2i-2)\, n^{2i-2}}$$

$$+ \frac{\theta B_2}{2i\, n^{2i}} = \log n - \frac{1}{2n} + \gamma,$$

γ being the Euler-Mascheroni constant.

Substitution in the expression for πS_n gives the desired formula

$$\pi S_n = 2n \log 2n + 2n(\gamma - \log \pi)$$

$$+ \sum_{j=1}^{i-1} (-1)^j \frac{B^2_{2j}(2^{2j}-2)\pi^{2j}}{j\,(2j)!\, n^{2j-1}} + R - \frac{\theta B_{2i}}{2i\, n^{2i-1}},$$

where R has the value given above and $0 < \theta < 1$. Our formula is the same as Professor Watson's, except in the remainder term, which for fixed i tends rapidly toward zero as n increases, as does Professor Watson's.

McGill University, Montreal,
August 1934.

* de la Vallée-Poussin, *loc. cit.* p. 87.

XXXIII. *A Crystallographic and X-Ray Investigation of some Diphenylamine Derivatives.*

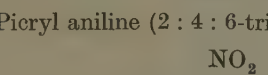
By R. G. Wood, *M.Sc.*, S. H. Ayliffe, *M.Sc.*, *and* N. M. Cullinane, *D.Sc.**

[Plate XIV.]

1. In a recent memoir [1] a preliminary investigation of the differently coloured crystalline modifications of certain derivatives of diphenylamine was reported. Certain physico-chemical properties, such as melting-points, densities, and solubilities, were compared, in order to ascertain if the forms were actually different in structure or identical apart from their appearance and colour. The conclusions therein arrived at have now been confirmed by crystallographic and X-ray methods.

Picryl Aniline.

2. Picryl aniline (2 : 4 : 6-trinitrodiphenylamine),



exists in two apparently different varieties, reddish-orange prisms and yellowish-orange needles.

In separating slowly from solution in ethyl acetate the compound was obtained as reddish-orange prisms showing no preferential direction of growth. When the amine was dissolved in a large volume of boiling alcohol, and a little concentrated hydrochloric acid added, yellowish-orange needles, up to about 20 mm. in length, were deposited. The comparison of the melting-points, solubilities, and densities of the two forms indicated that they were identical, and this conclusion has now been confirmed by a crystallographic examination. The most satisfactory confirmation, however, was obtained by the use of X-rays. Thus, the two modifications of picryl aniline, which were stated by Bamberger and Müller [2] to be chemical isomerides, have now been proved to be structurally identical, the difference in appearance and colour being due merely to a preferred direction of growth and different crystal texture.

* Communicated by Prof. R. T. Dunbar, Ph.D.

Microphotographs showing very clearly the differences in external appearance of the two varieties are reproduced in Pl. XIV, figs. 1 & 2.

3. *Goniometric Measurements.*

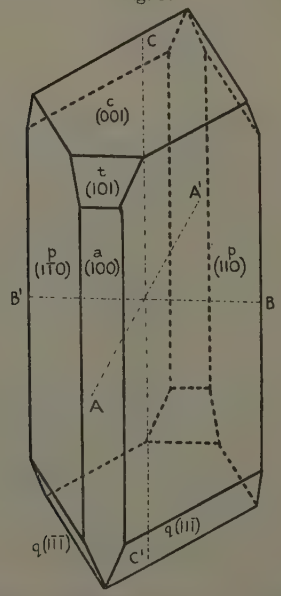
(*a*) Reddish-orange prisms.

Five specimens were examined, and the following table gives the mean angles between the various faces shown

Interfacial Angles on Reddish-orange Prisms of Picryl Aniline (fig. 3).

a (100)	*c* (001)	*a* (100)
53° 21′	75° 5′	26° 16′
p (110)	*p* (110)	*t* (101)
73° 18′	24° 24′	38° 7′
p ($\bar{1}$10)	*q* (11$\bar{1}$)	*c* (001)
53° 21′	80° 31′	115° 37′
a ($\bar{1}$00)	*c* (00$\bar{1}$)	*a* ($\bar{1}$00)

Fig. 3.



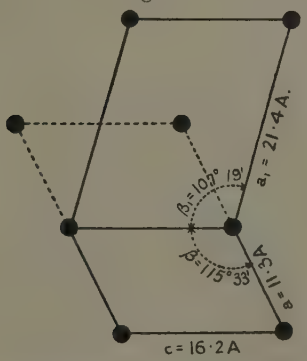
Typical crystal of reddish-orange form of picryl aniline.

in fig. 3. As is the case with all crystals described in this paper, the angles could not be measured with very great accuracy owing to surface imperfections. The mean value in most cases is correct to $\pm 5'$.

The crystal possesses a plane of symmetry and a two-fold rotation axis perpendicular to the plane. It therefore belongs to the monoclinic holohedral class. The b-axis coincides with the axis of symmetry, while the a- and c-axes must lie in the plane of symmetry inclined at an angle β to each other. On assigning indices to the faces, as shown in the above table, the axial ratios and angle are found to be

$$a : b : c = 1{\cdot}49 : 1 : 2{\cdot}12,$$
$$\beta = 115^\circ\ 33'.$$

Fig. 4.



Projection of unit cells of picryl aniline on the plane (010).

An X-ray and crystallographic investigation has been published by Hertel and Schneider [3], who examined a single variety of picryl aniline which they described as consisting of "scarlet-red needles," and which appears to be similar to the reddish-orange prisms described here. These investigators, however, chose a different a-axis, which we denote by a_1, leading to a different value for the axial angle :

$$a_1 : b : c = 2{\cdot}81 : 1 : 2{\cdot}12.$$
$$\beta_1 = 107^\circ\ 19'.$$

The relationship between the two sets of axes is shown in fig. 4.

(*b*) Yellowish-orange needles.

In this variety faces were found only around the zone of the needle-axis, and these were very poorly developed. There were three sets of parallel faces, the angular separations being approximately 73° 30′, 53° 15′, 53° 15′, suggesting that the zone-axis under observation was identical with the *c*-axis (zone *a p p a*) of fig. 3.

4. *X-ray Measurements.*

(*a*) Reddish-orange prisms.

Complete rotation photographs, using copper radiation, taken about the axes *a*, *b*, and *c*, gave the following results :—

$$\left.\begin{aligned} a &= 11{\cdot}3 \text{ Å.} \\ b &= 7{\cdot}62 \text{ Å.} \\ c &= 16{\cdot}2 \text{ Å.} \end{aligned}\right\} \; a : b : c = 1{\cdot}49 : 1 : 2{\cdot}13.$$

The density was determined by the flotation method (see § 8), the result being 1·570 gm./ml., confirming the value given by Cullinane, Embrey, and Davies[1]. From these data the number of molecules in the unit cell is found to be 3·95 (or 4).

Fig. 4 shows the projection of this cell and of that of Hertel and Schneider on the plane (010), the *b*-axis being normal to the plane and the same for both cells. It will be seen that Hertel and Schneider's cell (for which $a_1 = 21{\cdot}4$ Å., $b = 7{\cdot}62$ Å., $c = 16{\cdot}2$ Å., $\beta_1 = 107° 19'$) contains eight molecules and gives rise to a non-primitive lattice which is face-centred in the (010) planes.

(*b*) Yellowish-orange needles.

A rotation photograph with the needle-axis as axis of rotation gave a result identical with that given by the reddish-orange form with the *c*-axis as axis of rotation, except that the weaker spots were not obtained from the needle variety on account of the much smaller mass exposed. The value of the spacing in the direction of the needle-axis is 16·2 Å., which is identical with that obtained in the direction of the *c*-axis in the other form. Thus the X-ray measurements confirm the conclusion indicated by the goniometric measurements (§ 3 (*b*)).

For additional evidence Laue photographs were taken with the X-ray beam directed along the *c*-axis and the

needle-axis of the two varieties respectively. Apart from slight differences arising from a small and unavoidable inaccuracy in the setting of the needle, the two photographs were identical.

The density of the yellowish-orange form is 1·570 gm./ml., and is thus the same as that of the reddish-orange form.

There can, therefore, be no doubt that the yellow and red crystals of picryl aniline are identical in structure.

5. Besides the rotation photographs, a series of oscillation photographs was taken, each over a range of 15°, in order that the indices of the spots might be determined without ambiguity. For the analysis of the resulting diagrams the graphical method due to Bernal [4] was used. Indices of all general types (*hkl*) appeared (*i. e.*, all spectra for which *h*, *k*, and *l* all differ from zero), showing that the lattice is primitive.

The following features of special indices were found :—

(i.) $h0l$ absent when l is odd.

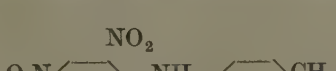
(ii.) $0k0$ absent when k is odd.

These characterize the space-group C_{2h}^5 ($P2_1/c$) in agreement with Hertel and Schneider. The other features of special indices found by these authors are accounted for by the non-primitive character of their lattice.

Four asymmetric elements are required for the description of this space-group, and so the four molecules in the unit cell may each be represented by an asymmetric element of the group.

Picryl p-*toluidine.*

6. Picryl *p*-toluidine (2 : 4 : 6-trinitro-4′-methyldiphenylamine),



has been obtained in two modifications, yellowish-orange and deep red. The latter is the stable variety, and is readily formed by crystallization from hot acetone. It consists of needles up to 5 mm. in length and about 1 mm. in thickness. The yellowish-orange crystals also consist of needles, and are of approximately the same

size as the red. They are best obtained by dissolving the amine in boiling carbon tetrachloride and allowing to cool. The comparison of the physico-chemical properties described by Cullinane, Embrey, and Davies[1] showed that the two varieties were not identical, and the further experiments now to be described confirm their conclusion.

Microphotographs of the crystals are given in Pl. XIV. figs. 5 & 6. The two photographs show the same crystals, some yellowish-orange and some red. Fig. 5 was taken in the yellowish-green light transmitted by a Wratten Series 6 safelight, and fig. 6 in the red light transmitted by a Wratten α-(contrast R) filter (no. 70). In fig. 5 there is a distinct contrast between the two varieties, the red crystals appearing black, due to absorption, while the yellowish-orange crystals transmit the light and appear transparent. In fig. 6 there is no contrast, but the surface details of the two varieties show equally clearly.

7. *Goniometric Measurements.*

(*a*) Yellowish-orange needles.

Only two zones were available for measurement, and the following table gives the mean results of measurements on five crystals. A drawing of the crystal is given in fig. 7.

Interfacial Angles on Yellowish-orange Needles of Picryl *p*-toluidine (fig. 7).

a (100)		*a* (100)	
	81° 51′		66° 33′
w (10$\bar{1}$)		*p* (110)	
	32° 17′		46° 54′
c (00$\bar{1}$)		*p* ($\bar{1}$10)	
	65° 52′		66° 33′
a ($\bar{1}$00)		*a* ($\bar{1}$00)	

The crystal class is monoclinic holohedral, the *b*-axis (axis of symmetry) being parallel to the axis of the needle. The indices to the faces are assigned as in the

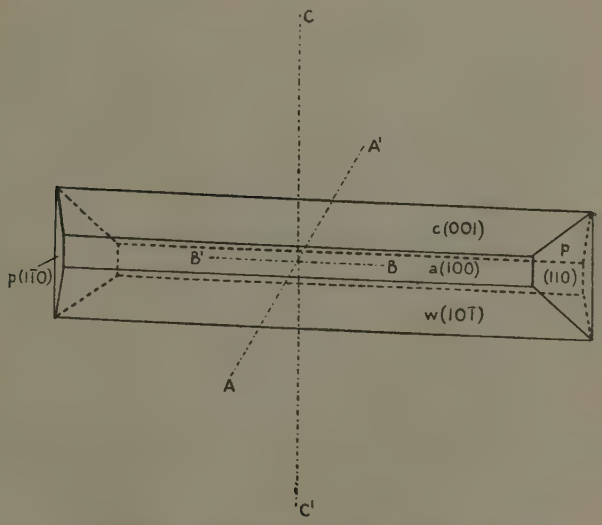
above table, and lead to the following values for the axial ratios and angle :—

$$a : b : c = 2{\cdot}53 : 1 : 2{\cdot}72,$$
$$\beta = 114° 8'.$$

(*b*) Red needles.

The faces were less perfect than those on the yellowish-orange crystals, and as no faces occurred around the symmetry axis it was only possible to obtain a rough

Fig. 7.



Typical crystal of yellow-orange form of picryl *p*-toluidine.

value of β from the goniometric measurements. Using this in conjunction with the X-ray rotation photographs, a more accurate value could be determined, and this is included below. The crystal is shown in fig. 8, and the following table gives the mean results of measurements on three crystals.

The crystal class is monoclinic holohedral, the *b*-axis (axis of symmetry) being at right angles to the needle-

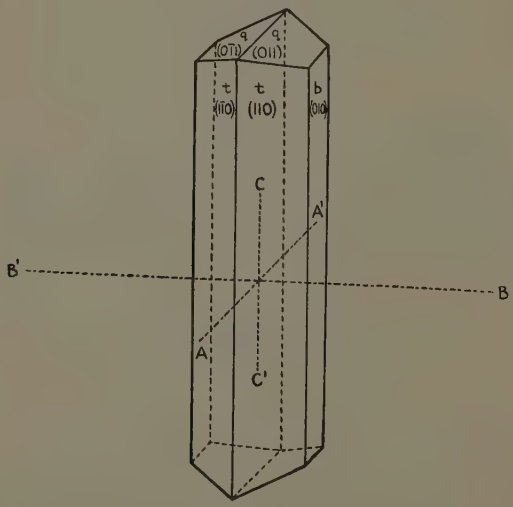
axis (c). The indices assigned as above lead to the following values for the axial ratios and angle :—

$$a : b : c = 0{\cdot}74 : 1 : 0{\cdot}37,$$
$$\beta = 103°.$$

Interfacial Angles on Red Needles of Picryl p-toluidine (fig. 8).

t	$(1\bar{1}0)$	q	$(0\bar{1}1)$
	71° 33′		39° 23′
t	(110)	q	(011)
	54° 14′		70° 19′
b	(010)	b	(010)

Fig. 8.



Typical crystal of red form of picryl p-toluidine.

A crystallographic examination of picryl p-toluidine has been reported by Lenk [5], who gives no details, but states that one modification consists of yellow or orange-yellow monoclinic needles, and the other of dark red needles belonging to the rhombic system. The above

results show, on the contrary, that both forms belong to the monoclinic system.

8. *X-ray Measurements.*

(*a*) Yellowish-orange needles.

Complete rotation photographs about the three axes gave the following results :—

$$\left.\begin{aligned} a &= 14{\cdot}9 \text{ Å.} \\ b &= 5{\cdot}94 \text{ Å.} \\ c &= 16{\cdot}3 \text{ Å.} \end{aligned}\right\} \; a : b : c = 2{\cdot}51 : 1 : 2{\cdot}74.$$

The flotation method was used to determine the density, as it was found that this gave more concordant results than the usual pyknometer method. A few well-formed crystals were immersed in carbon tetrachloride, which was then gradually diluted with alcohol so as to diminish the specific gravity of the liquid until the crystals tended neither to float on the surface nor to sink. The density of the mixture, in which the crystals are only very slightly soluble, was then measured with an accurate hydrometer. The density of the yellowish-orange needles was thus found to be 1·543 gm./ml.

These data lead to the value 3·9 (or 4) for the number of molecules per unit cell.

Two series of 15° oscillation photographs about the *a*- and *b*-axes were used to index the spots. The indices showed the same features as those for picryl aniline (see § 5). Hence the lattice is primitive and the space-group C_{2h}^5 ($P2_1/c$).

(*b*) Red needles.

The unit cell has the following dimensions :—

$$\left.\begin{aligned} a &= 12{\cdot}9 \text{ Å.} \\ b &= 17{\cdot}5 \text{ Å.} \\ c &= 6{\cdot}04 \text{ Å.} \end{aligned}\right\} \; a : b : c = 0{\cdot}737 : 1 : 0{\cdot}345.$$

The density was found to be 1·532 gm./ml., leading to 3·9 (or 4) molecules per unit cell.

The space-group was found, in the same way as before, to be C_{2h}^5 ($P2_1/c$).

9. *A Comparison of the Lattices found for Picryl aniline and the two Varieties of Picryl* p*-toluidine.*

The following table summarizes the measurements made on the above three crystals :—

	a.	*b*.	*c*.	*β*.	Density (gm./ml.).
	Å.	Å.	Å.		
Picryl aniline (both forms).	11·3	7·62	16·2 (needle-axis)	115° 33′	1·570
Picryl *p*-toluidine.					
Yellowish-orange needles.	14·9	5·94 (needle-axis)	16·3	114° 8′	1·543
Deep red needles.	12·9	17·5	6·04 (needle-axis)	103°	1·532

It will be seen that the two varieties of picryl *p*-toluidine are entirely distinct in structure. In the yellow form the needle-axis (*i. e.*, the preferred direction of growth) is the axis of symmetry (*b*), while in the red form it is the *c*-axis. In both cases the needle-axis coincides with the shortest side of the unit cell. Picryl aniline resembles the yellowish-orange form of picryl *p*-toluidine more closely, the *c*-spacings being almost identical. The needle-axis in picryl aniline, however, coincides with the longest side of the unit cell.

It is interesting to notice that the crystals of picryl aniline, which appear very different in the microphotographs (Pl. XIV.), are really identical, while the crystals of picryl *p*-toluidine, which appear more nearly alike, apart from colour, are entirely distinct in structure.

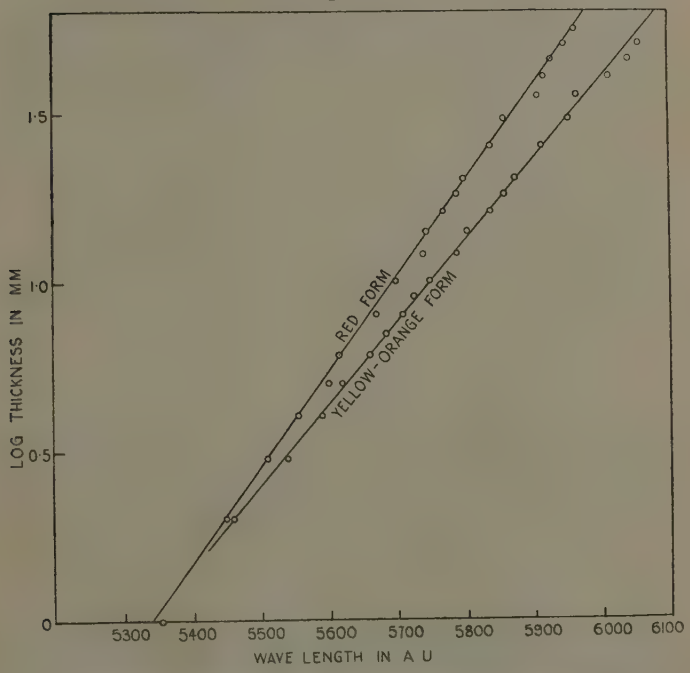
10. The above results show that the two varieties of picryl *p*-toluidine are different in crystal-structure. They are, moreover, different in molecular structure. This is shown by the chemical investigations of Cullinane, Embrey, and Davies [1] and by unpublished measurements on absorption spectra carried out by T. S. G. Jones. Their results may be summarized as follows :—

(*a*) The colours of equimolecular solutions of two forms in a given solvent are quite different, and these differences persist for many months.

(*b*) Different additive compounds with benzene are formed under the same conditions by the two modifications.

(*c*) For the absorption measurements 1·25 gm. of substance was dissolved in pyridine (purified by standing over sodium hydroxide and redistillation), the solution being made up to 25 c.c. The solution absorbs from the

Fig. 9.



Absorption curves for picryl *p*-toluidine solution in pyridine.

extreme ultra-violet to the green or orange, depending on the thickness of the absorbing layer. The wavelength corresponding to the edge of the band is plotted against the logarithm of the thickness in fig. 9, which shows that the curves for the two forms are quite distinct.

These results show that the two modifications are not dimorphs, but chemical isomerides. The type of isomerism is not known.

Measurements on the optical properties of the crystals are in progress, and it is hoped to publish the results shortly.

11. *Summary.*

Picryl aniline and picryl *p*-toluidine have been examined by crystallographic and X-ray methods, and the results confirm the chemical evidence as to the nature of the differences between the varieties of these substances.

The yellow and red crystals which picryl aniline yields are identical, differences being superficial only.

Picryl *p*-toluidine exists in two crystalline modifications which have different structures. The two forms differ in colour and other physical and chemical properties.

In conclusion we wish to thank Prof. R. T. Dunbar for granting us facilities for carrying out the work, and one of us (S. H. A.) is indebted to the University of Wales for a Dr. Samuel Williams post-graduate Studentship. Our thanks are also due to Mr. T. S. G. Jones, who carried out the measurements on absorption spectra.

University College, Cardiff.
Oct. 25th, 1934.

References.

(1) Cullinane, Embrey, and Davies, J. Phys. Chem. xxxvi. p. 1434 (1932).
(2) Bamberger and Müller, *Ber. d. d. chem. Ges.* xxxiii. p. 108 (1900).
(3) Hertel and Schneider, *Zeit. f. Phys. Chem.* xii. B, p. 109 (1931).
(4) Bernal, Proc. Roy. Soc. A, cxiii. p. 117 (1926).
(5) Busch and Pungs, *J. prakt. Chem.* lxxix. p. 548 (1909).

XXXIV. *On the Derivation of the Laws of Motion in the Theory of Relativity.* *By* Prof. Jan W. Weyssenhoff, *Dr. Phil., Stephen Bathory University, Wilno (Poland)**.

1. THE elementary treatment of kinematics in the relativity theory has been well developed, but the transition to dynamics presents still some difficulties. The method of introducing the fundamental equation of dynamics given by Lewis and Tolman† is rather suggestive and has found its way into many text-books. The

* Communicated by the Author.
† G. N. Lewis and R. C. Tolman, Phil. Mag. xviii. p. 510 (1909).

reasoning based on the ideal experiment with elastic spheres described by Lewis and Tolman is, however, somewhat complicated *. It can be simplified by considering inelastic collisions of two identical particles.

2. We shall deduce first the only formula of relativistic kinematics we shall need in the following development. Let us consider for this purpose two Galilean systems of reference S and S′, the latter moving with constant translation velocity **w** with respect to the former. The velocity components parallel and perpendicular to the relative velocity **w** in both systems are then connected by the well known formulæ :

$$v_{\parallel}=\frac{v'_{\parallel}+w}{1+\frac{wv'_{\parallel}}{c^2}}, \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$v_{\perp}=\frac{\sqrt{1-\frac{w^2}{c^2}}}{1+\frac{wv'_{\parallel}}{c^2}}v'_{\perp}. \quad . \quad . \quad . \quad . \quad . \quad (2)$$

From these equations we get

$$\sqrt{1-\frac{v^2}{c^2}}=\sqrt{1-\frac{v_{\parallel}^2+v_{\perp}^2}{c^2}}=\frac{\sqrt{1-\frac{w^2}{c^2}}\sqrt{1-\frac{v'^2}{c^2}}}{1+\frac{wv'_{\parallel}}{c^2}}, \quad . \quad . \quad . \quad (3)$$

and from equations (2) and (3)

$$v_{\perp}=\frac{\sqrt{1-\frac{v^2}{c^2}}}{\sqrt{1-\frac{v'^2}{c^2}}}v'_{\perp}. \quad . \quad . \quad . \quad . \quad . \quad (4)$$

This is the formula referred to above.

3. The problem is now this: assuming that it is possible to assign to each particle a vector **P** (called momentum) depending only on its velocity **v** (assumption A) in such a way that the sum of the momenta of all particles involved

* See, *e.g.*, M. Born, 'Einstein's Theory of Relativity,' transl. by H. L. Brose; G. Joos, 'Theoretical Physics,' transl. by J. M. Freeman (1934); also O. Halpern, "Relativitätsmechanik" in vol. v. of the *Handbuch der Physik* (Springer, Berlin, 1927).

in a collision remains constant in any collision and in any Galilean system of reference (assumption B), it is to be shown that :

$$\mathbf{P}=\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}\mathbf{v}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where m_0 is a constant for a given particle.

It follows from assumption A that the vectors **P** and **v** can be only parallel (or antiparallel), **P** vanishing with **v** ; we can write therefore

$$\mathbf{P}=m(v)\mathbf{v}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

$m(v)$ denoting a scalar function of the scalar velocity v. To find the form of this function let us consider any inelastic collision of two identical particles, their velocity before impact, measured in a Galilean system S, being denoted by **u** and **v**. After collision both particles move with a common velocity **w**. From the principle of conservation of momentum, whose validity has been assumed, it follows

$$m(u)\mathbf{u}+m(v)\mathbf{v}=\mathrm{M}(w)\mathbf{w}. \quad . \quad . \quad . \quad . \quad (7)$$

This equation contains two unknown functions. To get rid of the second one we shall consider only the velocity components perpendicular to **w**. The right-hand side of equation (7) vanishes, and we get

$$m(u).u_{\perp}+m(v).v_{\perp}=0. \quad . \quad . \quad . \quad . \quad (8)$$

Equations (7) and (8) must hold *mutatis mutandis* in any Galilean system of reference (assumption B). It will be sufficient, however, to consider, besides the system S, only such a Galilean system S′, in which both particles remain at rest after the collision. This system S′ moves with the uniform velocity **w** relative to S, so that the equations of section (2) may be applied without change of notations.

Now it follows from the fact that the particles come to rest in the system S′ that their velocities in this system before impact were equal and opposite :

$$\mathbf{u}'=-\mathbf{v}'. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

From equations (4), (8), and (9) we have

$$m(u).\sqrt{1-\frac{u^2}{c^2}}=m(v).\sqrt{1-\frac{v^2}{c^2}}. \quad . \quad . \quad (10)$$

The velocities u and v being independent of one another both sides of equation (9) must be constant. Taking into account equation (6) we get the formula (5). Q. E. D.

4. One point in the above argument needs perhaps some elucidation. It is at once apparent that two identical inelastic particles colliding with equal and opposite velocities come to rest, but the converse is perhaps not so obvious. If two identical particles come to rest in a collision their total momentum is seen to be zero; hence their momenta before impact are equal and opposite (and both particles approach one another on the same straight line). Does the equality of momenta imply the equality of the (scalar) velocities? One could imagine *à priori* the momentum of a particle assuming the same value for two different velocities. In this case the velocities of the colliding particles could be different. This is, however, impossible, since we can always find by means of equation (1) such a Galilean system S″ with respect to which the velocities before impact are equal and opposite. It follows that after the collision the particles would rest relative to the system S″; hence it cannot possibly differ from S′, and equation (9) must hold.

The above argument rests not only on the assumption A and B, but also on the form of the transformation-formula (1), showing that $v_{\parallel}$ is a monotonic function of w. We could avoid this reasoning altogether by introducing the additional assumption that v is a uniform function of P; such an assumption, however, seems less satisfactory from the physical point of view.

XXXV. *On the Theory of the Barkhausen-Kurz Oscillations.*

By Hannes Alfvén, *Uppsala, Sweden* *.

IN a recent paper † E. W. B. Gill gives a theory of the B.-K. oscillations on the basis of the space-charge effect found by him. According to this there exists a limiting current in the grid-plate space of a triode valve, and when this current is reached the conditions in the valve become unstable and oscillations start.

If we diminish the emission in a valve generating

* Communicated by Prof. J. S. Townsend, F.R.S.

† E. W. B. Gill, Phil. Mag. xviii. p. 832 (1934).

B.-K. oscillations with the frequency ω the oscillations stop at the emission I_0 ; but if we then measure the A.C. resistance R of the oscillator circuit at the frequency ω and $I<I_0$ we find * that it is

$$R=R_0\left(1-\frac{I}{I_0}\right),$$

where R_0 is the resistance with no emission in the valve. Consequently the examination of the valve in the pre-oscillating condition shows that there exists a phenomenon in the valve which decreases the apparent resistance of the circuit and that this decrease is proportional to the emission. In other words, the valve acts as a "*negative resistance*" *proportional to the emission*. If the emission is high enough this negative resistance compensates the original damping resistance of the circuit. If then the resulting A.C. resistance for the frequency ω is zero or negative, oscillations of this frequency start. Therefore it is very interesting to study the behaviour of a valve in the pre-oscillating condition, and it is likely that this will give the key to the whole B.-K. problem.

Such an investigation shows that at a constant emission the valve acts as a positive resistance at certain grid and plate voltages, and as a negative resistance at other voltages. We find negative resistances for several different ranges of the voltages, and it can be shown that *they correspond to the normal waves and the dwarf waves respectively*.

It seems impossible to explain these phenomena—which, in fact, are exactly those giving rise to the B.-K. oscillations—according to Gill's theory, for it is possible to observe them at such small emissions that the distortion of the field in the valve by the space-charge is negligible. On the other hand, it seems quite possible to explain in detail the behaviour of the valve according to the theory put forward in the cited paper †.

If we have constant tensions on the electrodes of a valve, and the emission is small, the electronic motion induces no current in the electrodes ; but if we apply an infinitely small A.C. potential between two of the electrodes (*e. g.*, between the grid and the plate) the

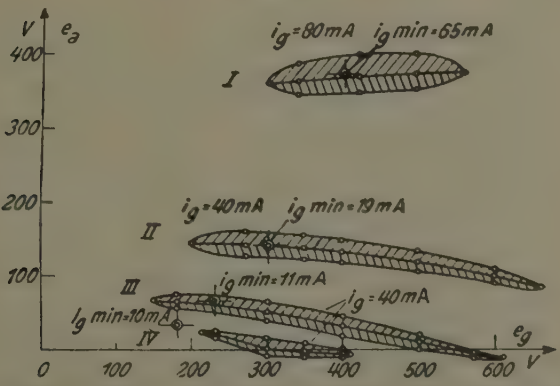
* Hannes Alfvén, "Investigations on the Ultra-short Electro-magnetic Waves," Uppsala Universitets Årskrift, 1934.

† Hannes Alfvén, *loc. cit.*

motion is disturbed, and this disturbed motion induces electrostatically an A.C. current in the electrodes (or, in other words, is equivalent to an A.C. current). A calculation of the relation between this current and the applied A.C. tension shows whether the valve acts as a positive or as a negative resistance.

The calculation is carried out under the assumption that the electronic motion is determined by the electrode tensions only, *i. e.*, the space-charge is considered to be negligible. We then find that the A.C. resistance of the valve is proportional to the emission, in accordance with the experimental results. Further, the variation of the resistance with the grid and plate tensions is expressed

Fig. 1.



Oscillation regions, measured experimentally by H. Collenbusch. (*Ann. d. Physik.* xiii. p. 201 (1932)).

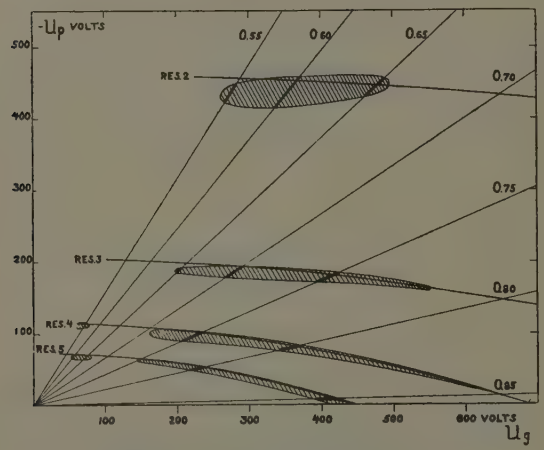
through rather complicated curves showing a good agreement with the experimental curves. This strongly supports the theory.

It must be pointed out that the theory explains the dwarf waves as well as the normal waves. It appears that the ratio between the $\lambda\sqrt{V}$-values of the normal waves and any of the dwarf waves need not be exactly an integer, but is often not very different from it.

I have tried to compare the theory with the experimental results of other authors. When a comparison is possible (often it is not because of the ill-defined conditions of many experiments) it seems to support the theory. As an

example the oscillation regions measured experimentally by H. Collenbusch * are shown in fig. 1, and the oscillation regions calculated according to my theory for a valve of the dimensions given by Collenbusch are shown in fig. 2. The small disagreements may be due to space-charge, of

Fig. 2.



Oscillation regions calculated theoretically.

which the theory does not take account, but on the whole it must be stated that the theory is able to give a good picture even of such complicated phenomena as the oscillation regions represent.

XXXVI. *Equilibrium Emission and Activity Changes in Oxide-coated Cathodes.* *By* A. J. Maddock, *M.Sc.*, *A.Inst.P.*†

Abstract.

Experimental results obtained on oxide-coated cathodes, subjected to various temperature conditions and not supplying emission current, are given in support of the

* H Collenbusch, *Ann. d. Physik*, xiii. p. 201 (1932).

† Communicated by L. J. Davies, M.A., B.Sc.

theory that a layer of barium exists at the outer surface of the oxide giving the enhanced emission. The existence of an equilibrium value of emission, and, therefore, of amount of surface covered with active material, for various temperatures is demonstrated. Data are also presented on ageing of such cathodes and the loss of activity that occurs on flashing at higher temperatures than normal.

AGREEMENT on the mechanism involved in the electronic emission from oxide-coated cathodes does not appear to have been finally reached, and the various theories proposed are still actively discussed. An account of some experiments carried out on the behaviour of such cathodes subjected to temperature treatments alone, *i. e.*, with no flow of emission current, are here put forward in support of the theory that a layer of barium * exists at the outer surface of the oxide giving the enhanced emission †, the activity of the cathode depending on the amount of surface covered: it has been shown by Becker ‡ that maximum activity occurs for a monatomic layer. When the active material has been formed within the body of the oxide subsequent heat treatment allows it to diffuse to the outside where the layer is formed, and, at any given temperature, equilibrium conditions are reached when the amount of material on the surface is constant and represents a balance between that diffusing outwards from the oxide coat and that lost by temperature evaporation. For the forming or ageing process it is often stated that the passage of electronic current through the coating is necessary, but some workers § hold that heat treatment alone will fully activate these cathodes; and our experience has always been that by simple high-temperature treatment a filament can be fully aged.

In the following account we present data which support

* As is usual barium is here intended to signify the alkaline earth metal or metals present in the coating.

† By this we mean that a layer of barium is necessary for emission; our experiments do not show whether the electrons are emitted directly from this layer or from the oxide surface immediately below.

‡ J. A. Becker, Phys. Rev. xxxiv. p. 1323 (1929).

§ *E. g.*, A. Gehrts, *Zeits. f. Techn. Phys.* xi. p. 246 (1930).

the view of the mechanism of oxide-coated cathodes as follows :—

(*a*) High-temperature treatment alone is sufficient for the ageing process ; cathodes after such treatment show high activity.

(*b*) A layer of barium exists at the outer oxide surface giving the enhanced emission, and, for any temperature, equilibrium is reached when the rate of loss of barium by temperature evaporation is equal to the rate of gain by diffusion from the body of the oxide ; the higher the temperature of the cathode the less the amount of surface covered.

Curves of activation and deactivation for particular filaments are shown for various temperatures, and the existence of the equilibrium point readily shown. In addition results are given of the loss of emission which occurs on flashing filaments at high temperatures.

Experimental Results.

All these experiments have been conducted on standard commercial valves with filaments having a mixture of barium and strontium oxides coated on a nickel core (" uncombined " cathodes). They were of the small L.F. type of battery valve, with a normal filament rating of 2 v., 0·15 a. Whilst oxide-coated cathodes do not give true saturation conditions, we shall nevertheless refer to the " total emission " (T.E.) as forming a convenient basis for comparison of the activity ; we have chosen an accelerating voltage of 150 for these measurements as the thermionic current-voltage curve is here on the sloping " saturation " part, well past the region of the 3/2 power relation. Emissions were measured with anode and grid tied together, and the accelerating voltage only applied for 1/100 sec. during measurements of T.E. This method was devised by Mr. L. J. Davies of the Research Laboratory, B. T.-H. Co., and utilizes a spring switch and ballistic galvanometer, calibrated directly in terms of current ; it obviates several difficulties when experimenting with these small filaments ; the T.E. is measured before the filament has had time to warm up, due to the passage of the large emission current, or be cooled by the heavy electron evaporation ; it

eliminates the small changes of emission such as take place in the usual methods with hand-operated switches, and also electrolytic effects are minimized.

1. *Ageing or Forming Process.*

That the passage of current through the coating, with consequent supposed electrolysis, is not necessary during the ageing process of these cathodes is readily evident from the fact that we are continually able to age oxide-coated cathodes of all types to a high value of T.E. simply by overrunning at an increased temperature for a short time. Many hundreds of valves have been aged in various ways with simple high-temperature treatment, and all give high values of T.E. ; whether filaments are outgassed on the pumps or not, overrunning at practically double normal voltage for times up to two minutes will effectively age them. As an example we quote the result on 100 valves with 2 v. filaments which were not heated until they had been sealed off and the getter flashed ; after this the filaments had no emission until they had been run at 3·8 v. for $\frac{1}{2}$ min., when large values of T.E. were obtained, and it was then only necessary to run at the normal filament rating and pass a small anode current for ten minutes to clean up all gas. We are led to the belief that this initial ageing treatment forms all the active material within the body of the oxide which it is possible to do, for we have never been able to reage such filaments after they have begun to lose emission after long service on life—in fact, further deactivation occurs with any attempted reageing treatment.

2. *Activation and Deactivation.*

We consider now experiments made on activation and deactivation of these filaments and equilibrium conditions at various temperatures. By deactivation we mean subjecting the filament to a higher temperature than normal, so causing a decrease of emission and by activation the restoration of activity by a lower temperature treatment. It is to be noted that in all these experiments we are dealing solely with temperature effects on filaments already well aged and all gas cleaned up ; the only time when emission current was allowed to flow was during the time of measuring the T.E., *i. e.*,

for several isolated periods of 1/100 sec. each; that these periods did not in any way activate filaments was shown by flashing a filament to a low value of T.E. and then activating back to its original value, measuring the T.E. every minute; again flashing similarly reduced the T.E. to the same low value, and this time the T.E. was only measured once, and that at the end of an activation time equal to that previously carried out, when it was found that the filament had activated to exactly the same figure, viz., its original T.E.

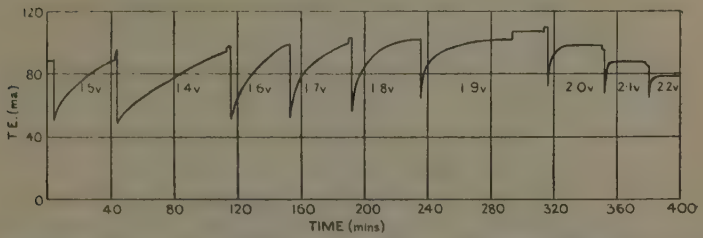
Throughout these experiments the filament voltage has been set at certain definite values as being more convenient to measure than temperature—in fact, it is not possible readily to determine the temperature, as we are dealing with commercial valves with the bulb covered with getter and the filament concealed by the other electrodes. However, the standard operating temperature of these filaments at their normal voltage (2 v.) is 800° C., and by assuming this to hold for every filament and that temperatures at other values of filament watts can be deduced directly from this, we obtain the following approximate temperatures, corresponding to the several values of filament voltage: 1·4 v., 650° C.; 1·5 v., 680° C.; 1·6 v., 710° C.; 1·7 v., 730° C.; 1·8 v., 760° C.; 1·9 v., 780° C.; 2·0 v., 800° C.; 2·2 v., 840° C.; 2·5 v., 900° C.; 3·0 v., 990° C.; 4·0 v., 1170° C.; 5·0 v., 1310° C.; which values we believe to be a good approximation, for taking the emission at these voltages for a filament in a well aged condition and with the above values for the temperatures we obtained good agreement for the constants in Richardson's equation, $i = AT^2 e^{-b/T}$, with other workers' carefully determined data; we get $A = 0{\cdot}04_4$ amps. deg.$^{-2}$ cm.$^{-2}$, $b = 1{\cdot}1_5 \times 10^4$° K., or $\phi = 0{\cdot}99$ volts. Espe * gives $A = 0{\cdot}0003$ amp. deg.$^{-2}$ cm.$^{-2}$, $b = 1{\cdot}15 \times 10^4$° K., $\phi = 0{\cdot}99$ v.

Experiments on activation at various temperatures were carried out on valve no. 2916; the filament was flashed at 3 v. for 5 sec. to deactivate it, and then activated at a low filament voltage; when the original emission had been regained the filament was again deactivated at 3 v. for 5 sec. and activation carried

* W. Espe, *Zeits. f. Techn. Phys.* x. p. 489 (1929). See also, *e. g.*, L. R. Koller, Phys. Rev. xxv. p. 795 (1925) and others.

out at a different temperature, and so on ; in cases where the original T.E. could not be regained activation was continued until the emission had reached a steady value. During each activation run the emission was measured at intervals with the filament voltage the same as that for activation ; by means of a subsidiary experiment, where the relation between the T.E. at various filament voltages to that at 2 volts was obtained, the values of T.E. obtained at these several temperatures could all be corrected to a common basis, viz., that corresponding to 2 v. (about 800° C.). Fig. 1 shows the results, being values of T.E. against time of activating at the voltages indicated ; each break in the curve denotes a deactivation carried out at 3 v. for 5 sec.

Fig. 1.



Activation curves at various temperatures (filament voltages as shown). T.E.'s all corrected to basis of 2 v. across filament. Breaks indicate deactivation flash at 3 v. for 5 sec. Valve 2916.

It will be observed that * :

(i.) The rate of activation increases with increase of temperature.

(ii.) The rate of activation decreases with increase of activity.

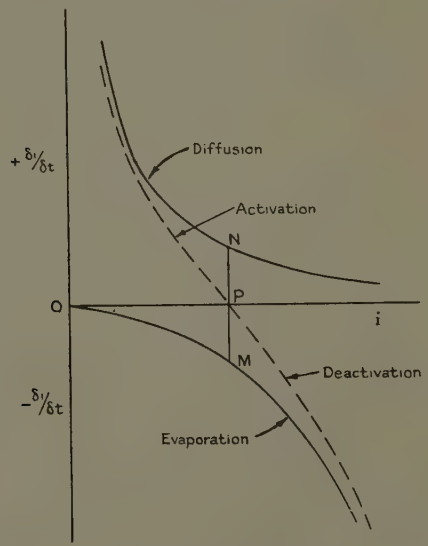
(iii.) It is possible to activate the filament to its original T.E. at voltages up to 1·9 v. (about 780° C.), but not

* Although we endeavoured to stop activation when the original T.E. had been reached, there has been a rise of 20 ma. throughout the course of the experiment ; this is explained by (*a*) on changing to a different temperature the filament changed its state a little immediately, probably by increased activation, and (*b*) some small errors would occur in the subsidiary experiment for correlating the T.E. at any filament voltage to that at 2 v. due to slight activation that would occur during the course of taking the measurements.

above. However, in these latter cases a steady state was reached for each temperature.

(iv.) At the lower activation temperatures the T.E. is still increasing when the filament had been brought back to its original emission, *i. e.*, it would, if allowed, have proceeded to a more active state the lower the temperature.

Fig. 2.



Supposed curves of rate of change of emission for one particular temperature.

Deactivation runs were also carried out on this filament by flashing at higher temperatures than normal for successive 1 sec. intervals, the T.E. being measured as rapidly as possible, to prevent activation, at 2 v. after each flash; flashing was continued until the T.E. became steady. The filament was then activated to a high value of T.E. and deactivation carried out at different temperatures, viz., 3, 4, 5 v. (about 990°, 1170°, 1310° C. respectively).

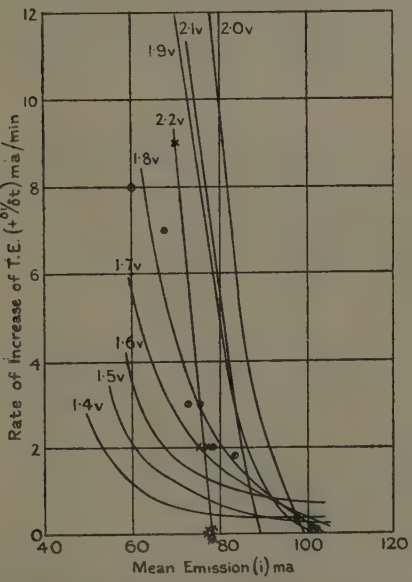
These showed that :

(i.) The rate of deactivation increases with increase of temperature.

(ii.) The rate of deactivation increases with increase of activity.

(iii.) The higher the deactivation temperature the lower the final value of T.E. reached.

Fig. 3.

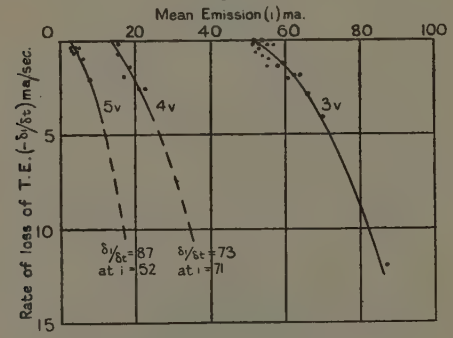


Rate of activation. T.E.'s all corrected to basis of 2 v. across filament. Valve 2916.

The above results support the theory that at the outer surface of the oxide there is a layer of barium which is continually being lost by temperature evaporation and replenished by outward diffusion from the body of the oxide, until finally, for any one temperature, a balance is obtained ; if the temperature is then raised the evaporation at first predominates and we get deactivation, whilst if the temperature were decreased activation would ensue, since the loss would decrease and more

material be retained on the surface. In general of course any experimental results we obtain are the resultant effect of the two processes, diffusion and evaporation. If the actual curves of rate of change of emission plotted against emission for one particular temperature were as sketched in fig. 2 then the curve we could obtain experimentally would be that dotted, the sum of the two full-line curves; this shows a definite equilibrium point P (where PN=PM) where there is no change in emission, activation occurring to the left of P and deactivation to the right. We see that the resultant curve obtainable has a point of inflexion at P, and is not of zero gradient at this point, and this is borne out by actual experiment.

Fig. 4.



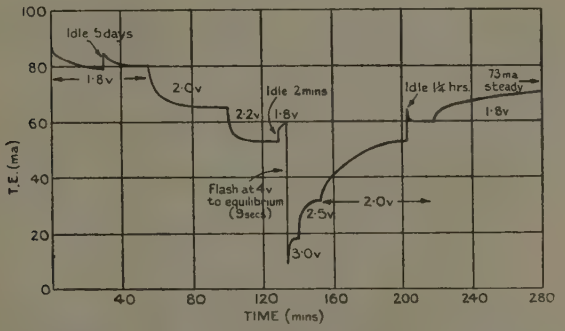
Rate of deactivation. T.E.'s all measured with 2 v. across filament. Valve 2916.

In practice $\delta i/\delta t$ has been taken as the change in emission over 1 min. of an activation run, and the corresponding T.E. (*i.*) taken as the mean of that at the beginning and end of this interval, and in fig. 3 results for activation runs at various temperatures on valve 2916 are plotted; to avoid confusion the actual points obtained are only shown for two curves, the others being mean curves through similar sets of points. The equilibrium T.E. is well shown, especially at the higher temperatures, and, as would be expected, this point recedes to the left as the temperature is raised, *i. e.*, the higher the activating temperature the lower the equilibrium value of T.E., or, in other words, the less active material there is present

on the oxide surface due to the increasing effect of evaporation. Similar curves showing deactivation at high temperatures on this valve are also given ($-\delta i/\delta t$ in ma./sec.) in fig. 4, and show that even at these very high temperatures there still exists the equilibrium T.E.

To obtain deactivation curves at low temperatures we must activate the filament at some very low rating and then note the changes that occur on running at the desired temperatures. This was done with the filament of valve no. 2863, which was activated to a high value of emission at 1·5 v. for a long period. Running first at 1·8 v. to equilibrium, then at 2·0 and 2·2 v., we obtain deactivation curves for these temperatures (see fig. 5).

Fig. 5.

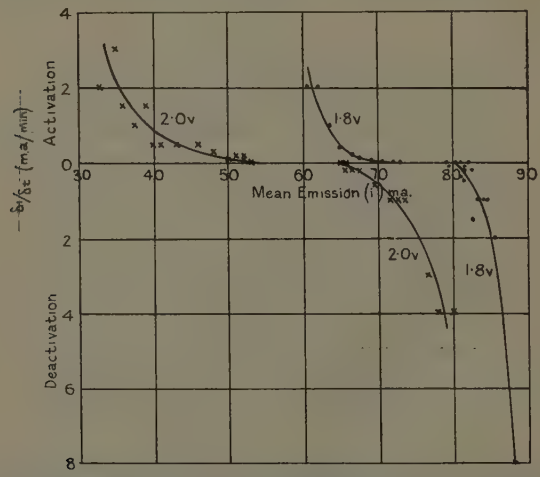


Activation and deactivation curves at various temperatures. T.E.'s all measured at 1·8 v. across filament. Valve 2863.

Flashing then at 4 v. reduced the T.E. to a low value. and thereafter activation runs were carried out at 3·0, 2·5, 2·0, and 1·8 v., and it is seen that the steady values of T.E. obtained at 2·0 v., after the activation and deactivation treatments, are approximately the same, as also are the corresponding values for 1·8 v. ; some small changes nearly always occur in oxide-coated filaments during the course of an experiment, and if the original T.E. can be regained within a few ma. it can be considered satisfactory. With this in view we can say that in the above experiments the filament had returned to its original state at the 2·0 and 1·8 v. treatments respectively. This can be shown very well by plotting, as

before, $\delta i/\delta t$ against i for these two temperatures for both activation and deactivation, and fig. 6 shows the results indicating an equilibrium value for the 2·0 v. runs between 55 and 65 ma. and for 1·8 v. between 73 and 80 ma. (T.E.'s in both cases being measured at 1·8 v.). The total emission is only a convenient means of estimating the state of the emitting surface, but since the emissions were all measured at the same temperature they do give a direct comparison of this state, and these results also demonstrate that the surface is more completely

Fig. 6.



Rate of change of T.E. Showing equilibrium values of emission. T.E.'s all measured with 1·8 v. across filament. Valve 2863.

covered with active material the lower the operating temperature. Becker * has shown that maximum activity is obtained for a monatomic layer, and to obtain such a layer, he vapourized barium on to the filament from an external source; glowing the filament after such deposition always lowered the activity below the maximum. It is interesting to note that for temperature

* J. A. Becker, Phys. Rev. xxxiv. p. 1323 (1929); Phys. Rev. xxxviii. p. 2193 (1931).

treatment alone of the filament we have never been able to pass through such a maximum, and therefore presumably the surface is never completely covered with barium.

3. *Loss of Emission by Deactivation.*

We have found that the loss of emission which occurs on deactivation by flashing, bears a simple relation to the value of the emission prior to flashing. If we flash a filament to deactivate it, activate at normal temperature to some higher value of T.E., deactivate again by the same flash, activate to any value, and so on, and plot the loss of T.E. (i_L) brought about by the flashing against the value before flashing (i), we generally obtain a set of points distributed on a straight line as in fig. 7, so that the loss of T.E. is directly proportional to the T.E. before flashing, and follows a relation of the form

$$i_L = k(i - i_0), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where i_0 and k are constants, the former being the value of i when $i_L = 0$, *i. e.*, the value of T.E. to which the filament can be flashed, and beyond which there is no further loss of T.E., for that particular flashing treatment, and k is the slope of the line.

From (1) we have

$$i_L/i = k(1 - i_0/i). \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Now $100\,i_L/i$ is the percentage loss which occurs on flashing, and equation (2) shows that as i decreases the percentage loss will decrease also unless $i_0 = 0$, when the percentage loss will be constant; thus any filament which shows a constant percentage loss means that $i_0 = 0$, *i. e.*, it indicates that the active material could be completely removed from the surface, whereas filaments showing a decrease in the percentage loss means that $i_0 \neq 0$ or that there is some value of emission (i_0) beyond which no further loss occurs for the particular flash considered. Whether these are true physical realities we cannot definitely say, as in any experiment the points get closer and closer together at the lower values of T.E., so that it is difficult to reach actual zero loss; but the results given in fig. 7 are approaching such a state, and when we notice the large range over which the relation holds

Fig. 7 (*a*).

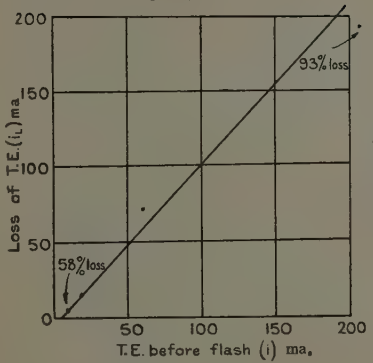


Fig. 7 (*b*).

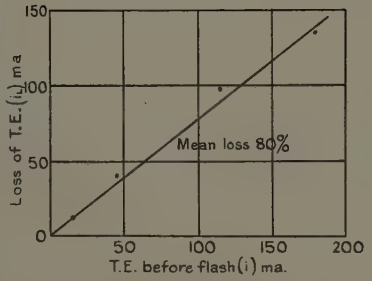
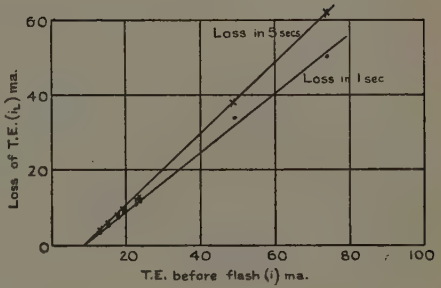


Fig. 7 (*c*).

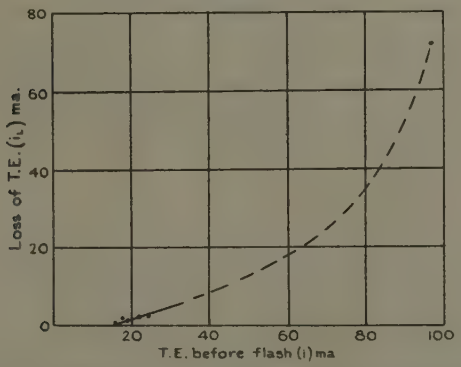


Loss of T.E. in flash and activation treatments.

Fig. 7 (*a*).—Valve 2918. Flashing at 4 v. for 1 min. and activating. $i_0 = 5$ ma. T.E.'s measured with 2 v. across filament.
Fig. 7 (*b*).—Valve 2864. Flashing at 4 v. for 1 min. and activating. $i_0 = 0$. T.E.'s measured with 2 v. across filament.
Fig. 7 (*c*).—Valve 2863. Flashing at 4 v. for successive seconds and activating. $i_0 = 8{\cdot}5$ ma. T.E.'s measured with 1·8 v. across filament.

extrapolation and its inference seem justified. In fig. 7 *a* the case of decreasing percentage loss over a large range (the end experimental points have losses of 93 and 58 per cent.) is given, showing a definite value for i_0, and in fig. 7 *b* the case when $i_0=0$ and the percentage loss is constant (mean of experimental values 80 per cent., whilst that calculated from equation (1) is 78 per cent.), whilst fig. 7 *c* shows results where flashing was split up into 1 sec. intervals, so that the loss in the first second or the total loss after a number of seconds could be plotted; as is to be expected, this only alters the slope of the line, but of course the value of i_0 is the same.

Fig. 8.



Loss of T.E. on flashing at 4 v. for successive seconds. No intermediate activation. T.E.'s measured with 2 v. across filament.

The preceding paragraphs of this section are concerned with flashing and some activation (for a long or short period is immaterial) between flashes. It might be expected that a similar relation to equation (1) would hold for repeated flashing with no intermediate activation. During the first second of flashing a very great loss of T.E. occurs, and the experimental results obtained show that this cannot fit a linear relation with the other points (see fig. 8), but that, considering the losses after the first second of flashing, a linear relation is indicated, though not so definitely as in the flash and activation treatment.

4. *Increase of Emission after a Cathode has been Inoperative.*

It has been observed in many cases that when a valve is left standing idle, even if only for a few minutes, the emission has increased by sometimes as much as 10 per cent. ; examples of such changes are shown in fig. 5, where, for certain reasons, the valve was left idle, *i. e.*, the filament at room temperature, and on testing after such periods at the same temperature as previously the emission had increased as shown. This is not a spurious phenomenon, as it occurs frequently, but as yet no satisfactory explanation has been found.

ACKNOWLEDGMENTS.

The author wishes to thank the B. T.-H. Co. (in whose laboratories this work was carried out) for permission to publish this paper, and feels particularly grateful to Mr. L. J. Davies for his continued interest in the work and for the many valuable and instructive discussions we have had on the subject.

XXXVII. *The Expanding Universe—an Alternative View.*

By J. A. CHALMERS, *M.A., Ph.D., Lecturer in Physics, Durham University (Durham Division), and* B. CHALMERS, *Ph.D., Lecturer in Mathematics and Physics, Sir John Cass Technical Institute, London, E.C.**

1. *Reasons against Interpreting the Hubble Effect as a Doppler Effect.*

SPECTROSCOPIC observations of the light emitted by the most distant astronomical bodies, namely, the spiral nebulæ, show that the spectral lines are shifted towards the red end of the spectrum, as compared with the spectral lines emitted in similar atomic processes in terrestrial sources. When the distances of the nebulæ concerned are measured by means of the Cepheid variables, it is found that the shift of the spectral lines is proportional to the remoteness of the nebula. This is known as the Hubble effect.

* Communicated by the Authors.

The Hubble effect is most simply interpreted as a Doppler effect caused by the nebulæ having large line-of-sight velocities of recession, the velocity of each nebula being proportional to its distance away. Eddington [1] interprets this as an expansion of the universe as a whole, and shows how this view is in accord with the theory of relativity.

There is, however, one serious difficulty in relation to the theory of the expanding universe, namely, that it cannot allow sufficient time in the past, since the expansion began, for the processes of stellar evolution which appear to have been necessary for the universe, and, in particular, our galaxy, to have reached its present state. Other theories, such as that of Milne [2], in which the Hubble effect is still ascribed to a Doppler effect are still less adequate in regard to the time scale (*cf.* Eddington [3]). In order to overcome the difficulty regarding the time scale, it is necessary to assume that the evolution of the stars themselves, and of our galaxy, has been going on much longer than the expansion of the universe; or else we can discard the idea of the expanding universe and find some other interpretation for the Hubble effect. Hubble [4] has pointed out that measurements on the apparent distribution of nebulæ with distance seem to be in accord with a fairly even distribution of the nebulæ in space, provided that the corrections applied do not consider the Hubble effect as a Doppler effect; but the distribution appears of an unexpected character if we apply a correction for motion in the line of sight. While the observational data are as yet insufficient to settle this point, the results encourage the search for an explanation of the Hubble effect alternative to the Doppler effect. The present paper suggests such an explanation, though the present ideas are, perhaps, not so much a different theory from that of the expanding universe as a different way of regarding the same facts.

2. *Possible Interpretations of the Hubble Effect.*

The Hubble effect shows a difference between the wave-lengths, as received and measured, of the light emitted at the nebula in the past and that emitted here now in a similar process. The wave-length, λ, is connected with the energy, E, of the light-quantum concerned

by the law $E = \frac{hc}{\lambda}$, where h and c are Planck's constant and the velocity of light respectively. The processes of absorption and measurement of the light from the nebula and from the local source are exactly similar, so that the difference in the measured λ's must correspond to a difference in the energies of the quanta received. This energy difference might be attributed to a loss of energy en route (as has been suggested by Zwicky [5]), or as a difference in energy on emission, either real or due to relative motion of the sources. In order to decide which of these alternatives is to be adopted as the true interpretation of the Hubble effect, it is first necessary to be able to compare energy measurements carried out at different times; thus one requires some convention regarding the choice of the unit of energy. Since there will be various possible ways in which an energy unit can be determined, we see that different interpretations of the Hubble effect may be equally valid. At present we will be concerned with considering a real difference to exist in the emission energies; in other words, we will use a definition of the unit of energy such that the number of such units in the quantum emitted in some specified physical process is not always the same. If E for the quantum differs, then one or more of the quantities h, c, and λ must differ, and, as we shall see, the definitions of units that we use are such that c and λ cannot alter, so that it must be h that differs for the two emissions. It has been shown experimentally that h is a constant here now, but there is nothing except the Hubble effect which can possibly give information on the value of h in the remote past. We are thus interpreting the Hubble effect as an alteration of h as between emission in the nebula and emission here now.

3. *Definition of the Units.*

We must now consider the definition of our units, and see how these may be consistent with an alteration in h, without invalidating ordinary observations.

(a) *Length.*

The normal unit of length is a standard metre, but we cannot imagine this to be transported to be at the nebula

at a date in the remote past; however, the standard metre contains a definite number of atoms, so that the inter-atomic distance is a definite fraction of a metre; the inter-atomic distance in a solid under definite conditions is a problem of atomic mechanics, and can, ideally, be solved, giving a result in terms of the atomic unit of length, a length equal to the radius of the first Bohr orbit in the hydrogen atom; this unit will, at the same time, define all atomic and inter-atomic distances and be a definite fraction of a metre. In order to compare lengths here now with lengths at the nebula in the past, an atom forms the obvious link, and atomic sizes will be considered the same in both cases. Thus the atomic unit, $r = \frac{h^2}{4\pi^2 m_0 e^2}$, where e and m_0 are the charge and mass respectively of an electron, must always be the same fraction of the unit of length. The wave-length λ of the light emitted in a definite physical process is always some numerical multiple of 1/R, where R is the Rydberg constant, $\frac{2\pi^2 m_0 e^4}{ch^3}$. Thus λ is proportional to $\frac{ch^3}{2\pi^2 m_0 e^4}$, which is $\frac{4\pi r}{\alpha}$, where α is the fine-structure constant. Eddington [6] supposes α to have the value 137, and α is certainly a numerical constant, so that, since r is constant, λ is constant, *i. e.*, the wave-length of the light emitted in a definite process is always the same number of units of length.

(b) *Time.*

Turning now to the unit of time, we see that the second, though our normal unit, is quite unsuitable as a unit to be used over long periods of time, since it is defined by the local and temporary facts of our own planet's present motion. We must, instead, find some unit of time which depends on phenomena we can imagine to exist at the nebula as well as here; for this we can take some type of periodic motion under electrical or gravitational forces. Since our definition of the permanent unit of length involves electrical forces and quantum phenomena in the atom, it is obviously preferable to define our unit of time in terms of similar phenomena. The atomic unit of time is the same as the period of the

electron in the first Bohr orbit of hydrogen, namely, $t=\frac{h^3}{4\pi^2 m_0 e^4}$, which is equal to $\frac{2}{c}\times\frac{1}{\mathrm{R}}$ or $\frac{8\pi r}{c\alpha}$, so that the unit of time is proportional to $\frac{1}{c}$ times the unit of length, so that c must be a permanent constant. This is not a property of light, but it depends on the definitions of the units we have chosen.

(c) *Mass.*

Considering now the unit of mass, it is incorrect to regard mass as the "quantity of matter," and hence to consider that m_0 will be constant, being the mass of an electron. Instead, we must consider mass as relating force and acceleration. If we say that the same acceleration is always produced by the electric force between two electrons a certain distance apart, then we must have e^2/m_0 constant.

In order to compare phenomena in a nebula in the past with phenomena here now, it is necessary to make certain definitions regarding the units to be used. The definitions we have chosen, involving electrical forces and atoms, require that r, c, and e^2/m_0 remain constant. These, and also α, can be constant even if h is multiplied by an arbitrary factor β, provided that e^2 and m_0 are also multiplied by the same factor β. No contradiction therefore arises if h is now β times as great as it was in the nebula in the past.

4. *Variation of "h" in Time.*

The difference in wave-length, as found in the Hubble effect, is proportional to the distance away, and, therefore, since the velocity of light is constant, to the time interval since the emission. Thus the Hubble effect is accounted for by a continuous variation in h; the exact way in which the wave-length of light received varies with the distance for large distances is not known with accuracy owing to observational difficulties, so it is not possible to state definitely how we suppose h to vary with time. If we consider λ_1 and λ_2 as the wave-lengths of light received from present and nebular emissions respectively, E_1 and E_2 the respective energies, and h_1 and h_2 the

values of h now and in the past, we have for emission at present $E_1 = \frac{h_1 c}{\lambda_1}$, for emission in the past $E_2 = \frac{h_2 c}{\lambda_1}$, (since λ is the same for a similar emission), and for absorption of E_2 now $E_2 = \frac{h_1 c}{\lambda_2}$; this gives $\frac{E_1}{E_2} = \frac{h_1}{h_2} = \frac{\lambda_2}{\lambda_1}$ so that if the variation of λ with time is strictly linear, then $1/h$ varies directly with time. If, however, λ varies in an exponential manner, then so too does h. The Hubble effect gives the numerical result that the present rate of variation is such that λ would be doubled for a time change of 2×10^9 years; if λ varies directly with time, then h would have been half its present value 2×10^9 years ago, and in another 2×10^9 years h will become infinite. If the variation of λ is exponential, then h doubles itself every $1{\cdot}4 \times 10^9$ years. The latter alternative seems more reasonable than a rapid rise of h towards an infinite value, and we shall see later that we can give it some justification.

5. *The Problem of Stellar Evolution.*

The problem of the time required for stellar evolution is affected by the suggestions here put forward; since h has a continuous variation from zero at a negative infinity of time, there is no limit to the length of time available in the past. Though the alterations in h, e^2, and m_0 would seem to indicate a "slowing-up" of atomic processes as we go back in time, this applies only to processes involving electrical forces, whereas stellar evolution must involve gravitational forces to a considerable extent. Our discussion of units does not involve any question of the constant G, and it seems that it cannot do so, for it is a matter of experiment rather than theory that will decide how the ratio between electrical and gravitational forces varies with time. Our explanation of the Hubble effect gives e^2 increasing with time, so that electrical force between two electrons increases with time; if the gravitational force between two electrons increases less rapidly, remains constant, or even decreases, then gravitational phenomena will not show the same "slowing-up" on going into the past, and hence, in a long period of time, more gravitational phenomena can

occur than would be expected from analogy with electrical phenomena, such as the Hubble effect is on our interpretation. If we accept the suggested explanation of the Hubble effect, then the results of theories of stellar evolution indicate the rate of variation of the fundamental constant G that is not included in that explanation. It is also possible that some inconsistencies in the theory of stellar evolution might be resolved by the distinction between electrical and gravitational forces.

6. "*Entropy*" *and* "*Atomic*" *Time Units.*

We see that the Hubble effect can be correlated with the changes in h, e^2, and m_0. We now have to consider whether these changes can be related to anything else that alters as time goes on. It is known (Eddington [7]) that the total entropy of the universe is continually increasing as time goes on ; so we might suggest that h, e^2, and m_0 depend on the total entropy of the universe. Though we shall not consider it further here, we can mention the possibility that the total number of particles in the universe may also alter with time, and thus may be related to the variation in h.

In order to consider the relation between entropy and the Hubble effect, it is convenient to consider the whole matter from the point of view of a hypothetical observer outside the universe, having definitions of units not necessarily the same as those for an observer inside the universe. This unit of length will be dependent on the size of the universe as a whole; but, as Eddington [8] has pointed out, there must be a relation between atomic sizes and the size of the universe, so that our external observer will choose the same atomic unit of length as we have previously considered. As regards the unit of time, the external observer will not choose the " atomic " unit, but a unit which depends on the universe as a whole. The " atomic " unit of time is the time occupied in some definite atomic process, and an internal observer, as we have already seen, will take this unit of time to be the same whenever in the history of the universe it may occur, so that the light emitted in a definite process has frequency ν_1 on emission whenever the emission occurs. The definition of the " atomic " unit of time as the time occupied in a certain physical process is not in itself

sufficient to give a measurement of time, for we require to be able to compare intervals of time with the unit. It seems that the most suitable definition to allow this is to say that if the probability of any specified change occurring in a unit time is a, and the probability of it occurring in another interval is b, then we can define the length of the second interval as b/a units, provided a and b are small. So that in t units of time the probability of the change is at, and, if there are n similar possible changes that might occur, the number that do occur in time t is *nat*.

Eddington [7] has pointed out that there is a connexion between the progress of time and the increase of entropy. He only uses this to indicate how the positive direction of time can be physically distinguished; it is clear that in any local region the change of entropy with time will be irregular. But, for the whole universe, it seems that the statistical effects will be such that the change of entropy of the whole universe with time (as ordinarily measured) will be regular. Since the change of entropy of the whole universe is a much less local matter than a normally measured time interval, it seems reasonable to select this as a way in which the external observer could measure time intervals. Let us say that equal changes in the entropy of the universe take place in equal intervals of what we can call "entropy time," and let us consider the relation between this and ordinary "atomic time." Imagine a physical process involving an increase $\delta\phi_1$ in the entropy of the universe, there being a probability a_1 that this process will occur in a unit of atomic time. If there are n_1 similar such processes, then in an atomic time t the number taking place will be n_1a_1t, and the consequent change of entropy $n_1a_1\delta\phi_1t$. Considering all other processes involving a change of entropy, the total entropy change in the universe in time t is $\Sigma n_1a_1\delta\phi_1t$, which we shall call $\Delta\phi_1t$; this measures the interval of "entropy time," which corresponds to an atomic time t. For a short interval of time, when the n_1's and a_1's are constant, then the entropy time and atomic time are proportional, but over longer periods it is clear that $\Delta\phi_1 = \Sigma n_1a_1\delta\phi_1$ is dependent on the epoch; in the absence of knowledge as to how the n's vary, the quantity $\Delta\phi_1$ must be left as it stands. It can be seen that $\Sigma n_1a_1\delta\phi_1$ will be nearly

proportional to $\int_{\text{II}}^{\infty} (\Sigma n_1 a_1 \delta\phi_1) dt$, the total change of entropy yet to occur before the final maximum entropy is reached ; if n_1 varies exponentially with time, and if the ratios n_1, n_2, etc., are constant, then the proportionality will be exact, and we can call $\Delta\phi_1$ equal to $p(\phi_0 - \phi_1)$, where ϕ_1 is the present and ϕ_0 the ultimate maximum entropy. In general $\Delta\phi_1$ may not be exactly proportional to $(\phi_0 - \phi_1)$, but we can see that the relation is not far from proportionality. In the following we will assume the proportionality, especially as the corrections for non-proportionality are not serious and could easily be followed.

7. *Interpretation of the Hubble Effect.*

The measurements on the Hubble effect consist in the comparison, at the same time and place, of two wave-lengths ; this must necessarily be the same for all observers. Different observers may differ in their interpretation of the Hubble effect, but not in their measurements of it. Observations of the Hubble effect consist in measuring the wave-lengths of light absorbed from two similar processes—one in the remote past and the other immediately previous to the absorption. Let us consider the matter from the viewpoint of the external observer. Suppose the atomic process considered to consist of emission of light of frequency ν_1, taking a time t_1, ν_1 and t_1 being measured in terms of atomic units ; since the same atomic process occurs, ν_1 and t_1 will be the same, in atomic units, whenever the emission occurs. If the entropy of the universe is ϕ_1, then the change of entropy in the atomic time t_1 is

$$\Delta\phi_1 = pt_1(\phi_0 - \phi_1),$$

while if the entropy is ϕ_2,

$$\Delta\phi_2 = pt_1(\phi_0 - \phi_2).$$

In each case there will be $n = \nu_1 t_1$ vibrations emitted, so that the external observer will measured the frequencies, using "entropy time," as $\frac{n}{\Delta\phi_1}$ and $\frac{n}{\Delta\phi_2}$ respectively. Thus the observed frequencies will be in the ratio

$$(\phi_0 - \phi_2)/(\phi_0 - \phi_1).$$

Since ϕ_2 is greater than ϕ_1, being the entropy at a later epoch, it follows that the ratio is less than unity, *i. e.*, that, when absorbed, the light emitted at an earlier epoch has a less frequency, or greater wave-length, than the light emitted at a later epoch. The farther off is the source, the longer is the time interval between the two emissions and so the greater is the difference between ϕ_2 and ϕ_1, and hence the greater is the Hubble effect. The change of wave-length is thus proportional to $\phi_2-\phi_1$, and hence to the time interval in entropy units. This gives, as suggested before, a Hubble effect which is not quite proportional to the distance away.

Measurements of the Hubble effect show that in the equation

$$\Delta\phi_1=pt(\phi_0-\phi_1)$$

the constant p is 5×10^{-10} year^{-1} or 2×10^{-18} sec.$^{-1}$ approximately. If the external observer uses the present year as his unit of time, then in a time 2×10^9 years (using entropy time) the universe will reach its final maximum entropy. But, to the observer inside the universe, using atomic units, which get progressively smaller compared with entropy units, there will be an infinite number of units of time before the entropy reaches its maximum; $(\phi_0-\phi)$ will reach half its present value in 10^9 years of entropy time, taking the present year as the unit; using atomic time units there will be $1{\cdot}4\times10^9$ years, so that we can say that the universe has a "half-value period" of $1{\cdot}4\times10^9$ years.

We have here related the Hubble effect to a progressive decrease in $(\phi_0-\phi)$; previously we have seen that the Hubble effect can be accounted for by a progressive increase in h, so that we may well suggest that $h\propto\frac{1}{\phi_0-\phi}$; but it does not seem possible at present to find any other basis for this proportionality. We might note that, to the external observer, h remains constant, and this is actually assumed above in relating the observed frequencies.

8. *Conclusion.*

We have shown that the Hubble effect can be accounted for without interpreting it as a Doppler effect. If this were merely another way of viewing the same phenomena

it might not be of much value; but the present ideas seem able to overcome the difficulties in the expanding universe theory, namely, the questions of the time scale and Hubble's results on the distribution of nebulæ. Further work along these lines is needed to show whether the present ideas can give the numerical agreement that Eddington [9] has found from the theory of the expanding universe.

References.

(1) Eddington, 'The Expanding Universe,' p. 13.
(2) Milne, 'Nature,' cxxx. p. 9 (1932).
(3) Eddington, *ibid.* p. 86.
(4) Hubble, 'Nature,' cxxxiii. p. 787 (1934).
(5) Zwicky, Proc. Nat. Acad. Sci. xv. p. 773 (1929).
(6) Eddington, 'Nature,' cxxiv. p. 840 (1929).
(7) Eddington, 'The Nature of the Physical World,' p. 79.
(8) Eddington, 'The Expanding Universe,' p. 87.
(9) Eddington, 'Proc. Roy. Soc. A, cxxxiii, p. 605 (1931).

XXXVIII. *The Compressibilities and Expansion Coefficients of Gases at Low Pressures.* *By* J. B. M. Coppock, *Ph.D.**

Introduction.

In a previous paper [1] an expression was derived from the Beattie-Bridgeman equation of state from which the compressibilities and expansion coefficients of gases could be rapidly calculated from a knowledge of the appropriate virial coefficients derived from the Beattie constants. The defect of obtaining such data in this way is that the constants of the Beattie-Bridgeman equation are only known for a limited number of gases. The object of this paper is to show certain relations between the compressibilities and expansion coefficients of gases and extend the uses of the equations previously developed.

The Compressibilities of various Gases.

The compressibility of a gas is represented by the relation

$$1+\lambda=\frac{p_0v_0}{p_1v_1},$$

* Communicated by Prof. R. Whytlaw-Gray.

where p_0v_0 is the limiting value of pv at zero pressure and p_1v_1 the value at one atmosphere pressure. The value of $1+\lambda$ for a given gas at any temperature T was previously shown to be given by the equation

$$(1+\lambda)_T = 1 - \frac{\beta}{R^2T^2+\beta},$$

where $R=0{\cdot}08206$ and β is of the appropriate value.

β is related to the Beattie constants by the equation

$$\beta = RTB_0 - A_0 - \frac{Rc}{T^2},$$

A_0, B_0, and c being characteristic constants for each gas, and T the temperature in ° Kelvin.

It was shown in the case of nitrogen that the equation was valid over a wide temperature range. The values of $1+\lambda$ at 0° C. have since been calculated, and are given in Table I. :—

TABLE I.

$1+\lambda$ values at 0° C.

Gas.	Beattie-Bridgeman, calculated.	Modified equation.	Most probable value.
Helium	0·99942	0·99942	0·99948
Neon	0·99950	0·99951	0·99951
Argon	1·00095	1·00095	1·00099
Hydrogen.....	0·99946	0·99946	0·99939
Nitrogen.....	1·00051	1·00051	1·00044
Nitrogen *	1·00047	1·00047	1·00044
Oxygen.....	1·00101	1·00101	1·00094
Carbon dioxide ...	1·00682	1·00678	1·0067
Methane	1·00233	1·00232	—
Air.....	1·00063	1·00063	—

* Smith and Taylor's values of the virial coefficient.

The most probable experimental values are those enumerated by Wild [2], except in the case of carbon dioxide, where the value of Cawood and Patterson [3] is given. It will be seen that the divergence between

the values calculated from the simple equation and those calculated by Beattie and Bridgeman [4] is extremely small.

Relations between the Expansion Coefficients and Compressibilities of various Gases.

It is becoming increasingly prevalent simply to state $1+\lambda$ values for gases at various temperatures, and to give practically no details of the actual pv measurements, thus making it virtually impossible to calculate expansion coefficients by the normal methods. If values of $1+\lambda$ are known at various temperatures it is possible to calculate the value of the virial coefficient β at any particular temperature from the relation

$$\beta = -\frac{\lambda}{(1+\lambda)} R^2T^2,$$

which is derived directly from the expression given for values of $1+\lambda$ in terms of β. The error involved in this calculation is small and will be of the same order as in the calculation of $1+\lambda$ values from the modified Beattie-Bridgeman equation, which in the case of carbon dioxide (a typical compressible gas) is practically negligible.

The coefficient of expansion of a gas at constant pressure α_v may be derived from the equation

$$\alpha_v = \frac{\left[RT_1 + \frac{\beta_1}{RT_1}\right] - \left[RT_0 + \frac{\beta_0}{RT_0}\right]}{\left[RT_0 + \frac{\beta_0}{RT_0}\right][T_1 - T_0]},$$

where $p=1$ atmosphere, $T_0=273{\cdot}13°$ C., and β_1 the value at $T_1°$ Kelvin.

The coefficient of expansion at constant volume α_p at an initial pressure of one atmosphere is given by the equation

$$\alpha_p = \frac{\left[RT_1 + \frac{\beta_1}{V}\right] - \left[RT_0 + \frac{\beta_0}{V}\right]}{\left[RT_0 + \frac{\beta_0}{V}\right][T_1 - T_0]},$$

where V = constant = 22·4131 litres at N.T.P.

Substituting in these equations appropriate values of $\beta = -\frac{\lambda}{1+\lambda} R^2T^2$ we obtain the following relationships:

$$\alpha_v = \frac{T_1\left[\frac{1}{(1+\lambda)_1}\right] - T_0\left[\frac{1}{(1+\lambda)_0}\right]}{T_0\left[\frac{1}{(1+\lambda)_0}\right][T_1 - T_0]}$$

at a constant pressure of 760 mm., and

$$\alpha_p = \frac{T_1\left[1 - \frac{\lambda_1}{(1+\lambda)_1} \cdot \frac{RT_1}{V}\right] - T_0\left[\frac{1}{(1+\lambda)_0}\right]}{T_0\left[\frac{1}{(1+\lambda)_0}\right][T_1 - T_0]}$$

at an initial pressure of 760 mm., where $(1+\lambda)_1$, λ_1 are the values at $T_1°$ absolute and $(1+\lambda)_0$, λ_0 are the values at 273·13° absolute (0° C.).

TABLE II.

Gas.	$(1+\lambda)_0$.	$(1+\lambda)_1$.	$10^6\alpha_v$.		$10^6\alpha_p$.	
			Calc.	Expt.	Calc.	Expt.
Helium ..	0·99947	0·99963	3659·0	3659[6]	3660·9	3661[6]
Hydrogen.	0·99938	0·99949	3659·8	3660·1[6]	3662·4	3662[6]
Nitrogen..	1·00046	0·99980	3670·3	3670·5[7]	3671·6	3671·5[7]
Oxygen...	1·00099	1·00012	3673·1	3673[5]	3672·5	3672·5[5]

Thus, if values of $1+\lambda$ are known at various temperatures, then the expansion coefficients may be derived from the equations suggested. Table II. uses known data, and indicates the reliability of the method. The coefficients have been calculated at 760 mm. and over a temperature range 0° to 100° C.

In the case of helium, hydrogen, and oxygen $1+\lambda$ values are calculated from *pv* data in the International Critical Tables, and the values for nitrogen are as determined by Holborn and Otto [5]. The references to the experimental values are given in the table.

This method of calculation has been applied to data obtained by Cawood and Patterson (*loc. cit.*) for various other gases, and appears to be quite satisfactory; the

values so obtained agree with those calculated by Batuecas [8].

For carbon dioxide a value $10^6\alpha_v^{0-21}=3733$ is found, which compares favourably with the experimental figure $10^6\alpha_v^{0-21}=3735$ of Chappuis [7]. In the case of nitrous oxide $10^6\alpha_v^{0-21}=3735$, which should be greater than the value $10^6\alpha_v^{0-75}=3729$ found by Coppock and Whytlaw-Gray [9]. For dimethyl ether values calculated are $10^6\alpha_v^{0-21}=4011$ and $10^6\alpha_v^{0-40}=3984$. The experimental value $10^5\alpha_v^{0-60}=392$ [9] compares favourably for such a compressible gas with $10^5\alpha_v^{0-60}=395$ obtained by linear extrapolation of the two calculated values to the 0–60° C. range.

The most important use of the equations suggested is that they may be written in such a form that $1+\lambda$ values at any temperature may be reduced to 0° C. Hitherto there has been no satisfactory method of accomplishing this, and such a knowledge of $1+\lambda$ at 0° C. is necessary in the determination of molecular weights by the limiting density method. Similarly from the equations previously developed values of $1+\lambda$ at T° may be obtained in terms of $1+\lambda$ at 0° C., and the coefficient of expansion of the gas at 760 mm. from T_0 to T_1° Kelvin (0° C. to t° C.). Thus we may write

$$(1+\lambda)_0=\frac{T_0}{T_1}(1+\lambda)_1[\alpha_v(T_1-T_0)+1], \quad . \quad . \quad (1)$$

$$(1+\lambda)_T=\frac{T_1\,.\,(1+\lambda)_0}{T_0\,.\,[\alpha_v(T_1-T_0)+1]}. \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Applying equation (2) to nitrogen, for which $(1+\lambda)_0=1{\cdot}00046$ (calculated from pv data of Holborn and Otto) and $10^6\alpha_v^{0-100}$ at 760 mm.$=3670{\cdot}5$, we find $(1+\lambda)_{100}=0{\cdot}99978$, the experimental value being $0{\cdot}99980$ [5]. Similarly for oxygen, where the experimental value of $(1+\lambda)_0=1{\cdot}00099$ and $10^6\alpha_v^{0-100}=3673$, we calculate $(1+\lambda)_{100}=1{\cdot}00015$, the experimental figure being $1{\cdot}00012$; such agreement is satisfactory. Cawood and Patterson have determined experimentally values of $1+\lambda$ at 0°, 21°, and 31° C. for carbon dioxide. Assuming their value $(1+\lambda)_0=1{\cdot}0067$, and taking $10^6\alpha_v^{0-21}=3735{\cdot}3$ and $10^6\alpha_v^{0-31}=3733{\cdot}8$, calculated from the data of Chappuis [7], we find $(1+\lambda)_{21}=1{\cdot}0053$ and $(1+\lambda)_{31}=1{\cdot}0047$, the corre-

sponding experimental values of Cawood and Patterson being 1·0052 (9) and 1·0048 (7).

Values of $1+\lambda$ at 0° C. may be calculated with equal confidence, especially where extrapolation of values at higher temperatures to 0° C. is uncertain. In the case of sulphur dioxide Cawood and Patterson find $(1+\lambda)_{30}$

TABLE III.

The Rare Gases.

Gas.	$10^6\alpha_p$.	Initial pressure.	$10^6\alpha_v$.	Pressure.	Temp. interval.	Reference.
		mm.		mm.	° C.	
Helium	3660·0	993	3658·5	991	0–100	10
	3661·3	1000	—	—	0–100	11, 12
	3662·5	700	—	—	0–100	13
	3660·8	1000	3657·9	1000	0– 50	5
	3660·7	1000	—	—	0–100	
	3661·8	613	—	—	0–450	14
	3660·4	994	3657·9	994	0–100	15
	3661·2	974	—	—	0–100	16
	*3661·3	760	3659·1	760	0–100	1
Neon	3660·5	687	3660·4	687	0–100	17
	3661·3	1007	3660·2	1007	0–100	
	3662·3	1362·8	3660·1	994·5	0–100	18
	3662·0	1000	—	—	0–100	12
	*3662·8	760	3660·6	760	0–100	1
Argon.....	3668	517	—	—	0–100	19
	3671·2	1000	—	—	0–100	12
	3669·1	746	—	—	8– 32	20
	3675	1000	3676	1000	0–100	5
	*3671·7	760	3672·4	760	0–100	1
Krypton	3689·9	1000	3696·7	1000	0–100	21
	—	—	3691·6	862	0–100	
Xenon	3720·2	1000	3739·5	1000	0–100	44

$=1·01681$; the value of α_v^{0-30} may be obtained by use of the linear relationship between α_v and temperature values where α_v is known over two or more temperature ranges. From the data for sulphur dioxide subsequently tabulated [40, 41] $10^6\alpha_v^{0-30}$ at 760 mm. is calculated to be 3955. Substituting these figures in equation (2) we find $(1+\lambda)_0=1·0249$, which is in exact agreement with the figure obtained by Cawood and Patterson by extrapolation

of the plot log A against log T $\left(\text{where } 1+\lambda=\frac{1}{1-A}\right)$ to 0° C., thus confirming the exactitude of their empirical relationship.

Collected Data on the Expansion Coefficients of Gases.

The author believes that no suitable collection of data exists from which expansion coefficients may be easily

TABLE IV.

Hydrogen and Nitrogen.

Gas.	$10^6\,\alpha_p$.	Initial pressure.	$10^6\,\alpha_v$.	Pressure.	Temp. interval.	Reference.
		mm.		mm.	° C.	
Hydrogen	3662·5	1000	3659·9	1000	0–100	22
	3662·7	1000	—	—	0–100	23, 24
	3662·6	700	—	—	0–100	13
	3662	623	—	—	0–450	14
	3662·0	727	3659·3	727	0–100	15
	3662·3	1095	3659·0	1095	0–100	6
	—	—	3660·9	760	0–100	25
	*3662·7	760	3660·3	760	0–100	1
Nitrogen	3670	744	—	—	0–100	26
	3668·9	620	—	—	0–450	14
	—	—	3670·2	760	0–443	27
	3675·4	1002	3677·0	1002	0– 20	22
	3675·2	1002	3675·0	1002	0– 40	
	3674·4	1002	3673·2	1002	0–100	
	3671·4	760	—	—	0–100	28
	3674·0	994	3673·4	994	0–100	15
	3669·6	654	—	—	0–100	29
	3674·4	1000	—	—	0–100	12
	*3672	760	3671	760	0–100	30
	*3671·8	760	3670·9	760	0–100	1

obtained over various temperature ranges and at low pressures. Tables have been constructed containing the majority of values which have been obtained experimentally; calculated values are marked with an asterisk. In each table the first member deviates least from the perfect gas and the order is according to increasing divergency. Simple equations are given in Table VI. for calculating α_v or α_p at any required pressure where experimental data is in existence for assigning values

to the constants. If this is not possible assume the limiting value of α_v or α_p at zero pressure to be 0·003661 and graphically interpolate values of the coefficients at any required pressure by assuming a linear relationship between the coefficients and pressure.

TABLE V.

Gas.	$10^6\,\alpha_p$.	Initial pressure.	$10^6\,\alpha_v$.	Pressure.	Temp. interval.	Reference.
		mm.		mm.	°C.	
Carbon monoxide.	3666·7	760	3668·8	760	0–100	34
	3671	760	—	—	0–250	35
	—	—	3674	760	11– 48	9
	*3673	760	3672	760	0–100	30
Air	3674·4	1000	3672·8	1000	0–100	22
	3675	1000	3676	1000	0– 50	32
	3675	1000	3674	1000	0–100	33
	—	—	3672	760	25– 50	9
	*3671·6	760	3671·1	760	0–100	1
Oxygen	3674·2	1000	—	—	0–100	12
	3674·4	708	—	—	0–100	31
	3677·0	1000	3679·2	1000	0– 50	5
	3675·7	1000	3676·3	1000	0–100	
	*3672	760	3673	760	0–100	30
	*3673·5	760	3674·6	760	0–100	1
Nitric oxide ...	—	—	3707·4	760	−140– 0	36
	—	—	3679	760	11– 48	9
Methane.....	3680	760	3683	760	0–100	37
	*3678	760	3681	760	0–100	30
Carbon dioxide.	—	—	3727	738	0– 32	25
	3726·3	1000	3741·0	998	0–100	7
	—	—	3753·6	998	0– 40	
	—	—	3760·3	998	0– 20	
	—	—	3725	760	11– 48	9
	*3712	760	3723	760	0–100	30
	*3710	760	3721·7	760	0–100	1
Nitrous oxide ..	—	—	3706·7	1000	0–100	38
	3676	760	3720	760	0–100	34
	—	—	3729	760	25– 50	9
	*3719	760	3732	760	0–100	30
Ethylene	—	—	3735	760	25– 50	9
	*3722	760	3735	760	0–100	30
Hydrogen chloride.	3681·2	760	—	—	0–100	34
	*3721	760	3734	760	0–100	30

TABLE V. (*cont.*).

Gas.	$10^6\,\alpha_p$.	Initial pressure.	$10^6\,\alpha_v$.	Pressure.	Temp. interval.	Reference.
Chlorine	3807	760	3833	760	0–100	39
	—	—	3873	760	0– 50	
	*380†	760	383†	760	0–100	30
Cyanogen	3682	760	3877	760	0–100	34
	383†	760	387†	760	0–100	30
Sulphur hexafluoride.	—	—	3808	760	11– 48	9
Sulphur dioxide.	3845	760	3903	760	0–100	34
	—	—	3900	760	0–100	40
	—	—	3963	760	0– 20	41
	—	—	3787	334	0– 22	
	*384†	760	388†	760	0–100	30
Ammonia	—	—	3802	740	0–100	42
	—	—	3854	760	0– 50	
	3802	991	3847	760	0–100	43
	3906	778	—	—	0– 20	
	*3767·8	760	3790	760	0–100	1
	*377†	760	380†	760	0–100	30

† $10^5\alpha_v$ or $10^5\alpha_p$.

Table VI. gives values of the constants in the following equations, from which values of α_v at a constant pressure P or α_p at an initial pressure P′ may be calculated for the temperature range stated.

$$\alpha_v^{0-t^\circ} = \alpha_v^{\text{limiting}} + \mathrm{KP},$$

$$\alpha_p^{0-t^\circ} = \alpha_p^{\text{limiting}} + \mathrm{K'P'},$$

where K, K′ are constants, $\alpha_v^{\text{limiting}}$, $\alpha_p^{\text{limiting}}$ the limiting values of the coefficients at zero pressure; P and P′ should be stated in metres.

A critical survey of individual results is largely unnecessary, but an examination of the tables given contributes to the view expressed by Coppock and Whytlaw-Gray [9] that some standard material be used for expansion coefficient measurements. Fused silica was suggested, since with this material adsorption effects were shown to be small and therefore least likely to produce abnormally high results. In the case of the more perfect

gases, such as hydrogen, helium, and nitrogen, comparable data is mainly in good agreement ; but even in such cases where values of the two coefficients are measured over

TABLE VI.

Gas.	$10^6K'$.	$10^6\alpha_{p}^{\text{limiting}}$.	10^6K.	$10^6\alpha_{v}^{\text{limiting}}$.	Temp. interval.	Referenc
					° C.	
Helium	−·4	3660·4	−1·9	3660·4	0–100	6
	—	—	−2·8	3660·7	0–100	15
Hydrogen	1·7	3660·4	−1·2	3660·4	0–100	6
Nitrogen	13·4	3660·4	12·7	3660·4	0–100	6
	13	3661	—	—	0–100	22, 28
	12	3661	—	—	0–100	26
Carbon monoxide.	—	—	14	3663	0 - 60	9
Air	—	—	13	3662	0– 75	9
Oxygen	13	3665	—	—	0–100	31
Nitric oxide ..	—	—	22	3662	0– 60	9
Krypton ...	29	3660·9	35·8	3660·9	0–100	21
Carbon dioxide.	58	3668	69	3671	0–100	
	—	—	91	3663	0– 40	7
	—	—	99	3661	0– 20	
	—	—	84	3661	0– 60	9
Xenon	59·3	3660·9	78·5	3660·9	0–100	44
Nitrous oxide.	—	—	89	3661	0– 75	9
Ethylene	—	—	96	3662	0– 75	9
Sulphur hexafluoride.	—	—	191	3663	0– 60	9
Dimethyl ether.	—	—	345	3658	0– 60	9

Note.—Values with reference number 9 are over the ranges 25–50° C. and 11–48° C. referred to the volume at 0° C. ; these have been shown to be equal to values obtained over the ranges 0–75° C. and 0–60° C.

a definite pressure range with subsequent extrapolation to zero pressure, with a view to obtaining the values of absolute zero, adsorption phenomena would probably produce a marked effect at pressures of about 300 mm.

This would occur if substances such as platinum-iridium were used for the construction of expansion vessels; such effect is particularly noticeable with more condensable gases, such as carbon dioxide, where limiting values of α_p^{0-100} are as high as 0·003671 [7], and is well marked in the case of glass [9]. Thus in the case of the condensable gases discrepancies between the values of various observers may be largely attributed to the different degrees of adsorption on the varieties of expansion vessels used. It would appear therefore most advantageous if some standard material were adopted for comparisons of the thermal expansion of gases.

References.

(1) Coppock, J. Phys. Chem. xxxvii. p. 995 (1933).
(2) Wild, Phil. Mag. xii. p. 41 (1931).
(3) Cawood and Patterson, J. Chem. Soc. p. 619 (1933).
(4) Beattie and Bridgeman, J. Amer. Chem. Soc. l. p. 3151 (1928).
(5) Holborn and Otto, *Zeit. f. Physik*, x. p. 367 (1922); xxxiii. p. 1 (1925).
(6) Henning and Heuse, *Zeit. f. Physik*, v. & vi. pp. 285, 264 (1921).
(7) Chappuis, *Trav. et Mém. du Bur. Int.* xiii. p. 93 (1907).
(8) Batuecas, *J. de Chimie Physique*, xxxi. p. 65 (1934).
(9) Coppock and Whytlaw-Gray, Proc. Roy. Soc. A, cxliii. p. 487 (1934).
(10) Heuse, *Zeit. fur Physik*, xxxvii. p. 157 (1926).
(11) Onnes, *Versl. Akad. Amsterdam*, p. 501 (1907)
(12) Cath and Onnes, *Comm. Leiden*, clvi. (1922).
(13) Travers, Senter, and Jacquerod, Phil. Trans. A, cv. (1903).
(14) Holborn and Henning, *Ann. der Phys.* xxxv. p. 761 (1911).
(15) Heuse and Otto, *Ann. der Phys.* ii. p. 1012 (1929).
(16) Keesom and Van der Horst. *Comm. Leiden*, clxxxviii. *a* (1928).
(17) Heuse, *Zeit. für Physik*, xxxvii. p. 157 (1926).
(18) Heuse, *Ann. der Phys.* iv. p. 778 (1930).
(19) Kuenen and Randall, Proc. Roy. Soc. A, lix. p. 60 (1895).
(20) Leduc, *Ann. de Physique*, ix. p. 5 (1918); *Compt. Rend.* clxiv. p. 1003 (1917).
(21) Heuse and Otto, *Physikal. Zeit.* xxxv. p. 57 (1934).
(22) Chappuis. *Trav. et Mém. du Bur. Int.* vi. p. 128 (1888).
(23) Onnes, *Comm. Leiden*, lx. (1901).
(24) Onnes and Boudin, *Versl. Akad. Amsterdam*, p. 224 (1900).
(25) Richards and Mark, *Zeit. für phys. Chem.* xliii. p. 475 (1903).
(26) Day and Clement, Amer. Journ. Science, xxvi. p. 405 (1908).
(27) Eumorfopoulus, Proc. Roy. Soc. A, xc. p. 189 (1914).
(28) Chappuis and Harker, *Trav. et Mém. du Bur. Int.* xii. p. 89 (1902).
(29) Keyes, Townshend, and Young, J. Maths. Phys. Mass. Inst. Tech. i. p. 243 (1922).
(30) Leduc, *Ann. de Physique*, v. p. 180 (1916).
(31) Makower and Noble, Proc. Roy. Soc. A, lxxii. p. 379 (1903).
(32) Holborn and Otto, *Zeit. für Physik*, xxii. p. 77 (1924).
(33) Holborn and Schultze, *Zeit. für Physik*, xlvii. p. 1089 (1915).
(34) Regnault, *Ann. Chim. et Phys.* iv. p. 5; v. p. 52 (1842).
(35) Amagat, *Compt. Rend.* lxxiii. p. 183 (1871).
(36) Adwentowsky, *Bull. Acad. Cracovie*, p. 742 (1909).

(37) Keyes, Smith and Joubert, J. Maths. Phys. Mass. Inst. Tech. i. p. 191 (1922).
(38) Jolly, Pogg. *Ann.* Jubelband. p. 82 (1874).
(39) Pier, *Zeit. für Phys. Chem.* lxii. p. 385 (1908).
(40) Walker, Proc. Roy. Soc. A, lxxii. p. 24 (1903).
(41) Leduc, *Compt. Rend.* cxvii. p. 219 (1893).
(42) Chappuis, Wied. *Ann.* viii. p. 1 (1879).
(43) Perman and Davies, Proc. Roy. Soc. A, lxxviii. p. 28 (1906).
(44) Heuse and Otto, *Physikal. Zeit.* xxxv. p. 628 (1934).

Chemistry Department,
Robert Gordon's Colleges,
Aberdeen.

XXXIX. *Practical Methods of determining the Coefficient of Photo-elastic Extinction.* *By* J. Kuno, *Dr. Eng., Assistant Professor of Civil Engineering, Kyushu Imperial University, Fukuoka* *.

[Plates XV. & XVI.]

I. Introduction.

THE photo-elastic constant is used for determining the magnitude of stresses from the isochromatic lines photographed with the specimen placed in the beam of polarized monochromatic light. In a substance such as phenolite under load, the optical creep is unavoidable. The photo-elastic constant for practical use is therefore to be determined under the condition that the creep is properly taken into account.

In the case of phenolite the photo-elastic constant is also desirable to be measured for every plate supplied, because this material is apt to be affected by the constituents and also by the temperature and time of baking †. Hence it follows that a simple and speedy method is needed in practice to determine the constant.

A great deal of labour and time are required when the method described previously by the author ‡ is adopted. Such tediousness prevents the method from satisfying the practical demand, notwithstanding that it gives the most general and correct values. In this paper other methods are studied in order to give a particular value of the constant with ease.

* Communicated by the Author.
† Phil. Mag. xvi. p. 361 (1933).
‡ Phil. Mag. xii. p. 508 (1931).

II. Necessary Conditions for Measurement.

1. *Inclusion of the Creep.*

In *the zero method* (or the colour-matching method) of photo-elastic observation the specimen is kept at a constant load for a long time. The point measured at the beginning includes less creep than the point measured at the end. The uniform creep is not ensured in the compensating specimen, too, as the stress in it is to be changed from time to time. It may therefore be said that this method is not preferable when the creep exists.

The photograph method, on the other hand, keeps the creep at various points uniform. In this case, however, special attention must be paid to the value of the photo-elastic constant. It is to be determined under the condition that the creep of t minutes is included, if the isochromatic lines of the specimen are photographed at the instant t minutes after loading. Experiments show that the value of t should be taken greater than zero, the reasons being as follows :—

1. It is not easy to apply the load suddenly enough without starting vibration and, in addition, to observe the optical effect at the same instant.

2. The creep is remarkably rapid immediately after loading. A slight lapse of time in measuring may lead to a considerable error in the optical effect observed.

The author takes t as one minute.

2. *Interruption of Load.*

The points of both higher and lower stresses are subjected to the same creep of t minutes, if the specimen is photographed at t minutes after loading. Then the photo-elastic constant must also satisfy the same condition. For this purpose a particular method of loading is needed for measuring the constant, namely :—

(1) The optical effect is measured at t minutes after loading under the load P_1. Then the load is removed at once.

(2) Keep the specimen unloaded for $2t$ minutes or more to avoid the effect of creep on unloading. The load P_2 is then applied, after which the optical effect

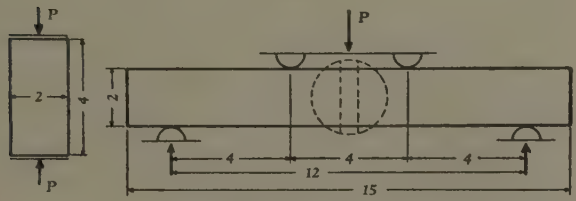
is measured at t minutes as before. Then the load is again removed. P_3, P_4, ... are applied in like manner.

This procedure we shall refer to as *the load interrupting method.* Another method of increasing the load step by step without interruption fails to give the constant for the photograph method.

3. *Increment Method.*

Initial stresses are unavoidable in phenolite. It has hitherto been put in practice to remove them by proper annealing. It is, however, apparent that to eliminate the effect of initial stresses has an advantage over removing them. This is achieved by adopting the method that increments of load and corresponding increments of the optical effect are measured. Whether initial

Fig. 1.



Specimens to determine the constant.

stresses exist or not becomes out of the question. The trouble of annealing also disappears.

III. Specimens.

It is well known that the compressive strength of concrete varies according to the shape and dimensions of the specimens tested. It may be that the same fact exists in phenolite. Accordingly it is reasonable to decide upon a standard form of the specimen. The shape of the specimen and the loading device should be as simple as possible.

The compression and the bending specimens to be used are now determined, the tension one being rejected. After consultation with Dr. Tuzi, the dimensions were chosen as shown in fig. 1, each being 5–6 mm. thick.

IV. Compression Method.

1. *Theory.*

The optical effect of a compression specimen is measured by Babinet's compensator, the specimen and the compensator being set properly in the beam of polarized green light (5461 Å.). Regarding the compensator, the notation a stands for the distance through which a wedge is moved in order to replace a dark band with an adjoining one. Inserting the specimen before the compensator, the band will displace. To put back the shifted band to its original position, the wedge is to be moved by the distance x. This distance is referred to as *the band displacement*, which is to be measured at t minutes after loading on every stage of the load. Then the order of extinction may be given by the equation $N=x/a$.

If the specimen is b cm. wide and d cm. thick, and is subjected to the load P kg., the stress is given by the equation

$$S=P/bd.$$

Inserting these into *the law of photo-elastic extinction* *, i. e., $S \cdot d=K_t \cdot N$, we have

$$P=(b/a)\,K_t x \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where K_t is the coefficient of extinction, the creep of t minutes being taken into account.

Plotting the load to the band displacement ($P-x$ diagram), the points plotted will nearly form a straight line, as shown by the equation (1). Denoting the tangent of the slope of the $P-x$ line by m, we have

$$K_t=(a/b)\,m \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

When the initial stress exists at the point observed, it may be supposed that a virtual load P_0 causes it, and that the total load $(P+P_0)$ gives the measured displacement x. Then we have, instead of (1),

$$P=(b/a)\,K_t x-P_0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

As P_0 may be considered constant at a point, the equation (3) is not different from (1) in the slope. The initial stress has no influence upon the value of K_t.

The equation (3) further suggests that a load applied constantly throughout the experiment may not be taken

* Phil. Mag. xvi. p. 354 (1933).

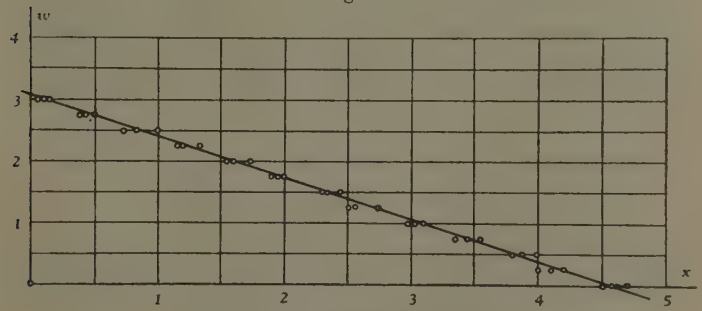
into account. Such a load as due to the weight of the lever and its accessories may be safely neglected. In other words, the increments only are needed.

2. *Method of Measurement.*

The load is applied to the specimen by means of a lever system, the magnification of which is five. Fig. 2 (Pl. XV.) shows a part of the loading device. To make measurement clear double crosses are marked on the surface of the specimen. The band displacements are observed on several stages of the load, the weight w at the lever end being varied from zero to 3 kg. On each stage of the load three operations are carried out, namely :—

1. To apply the load to the specimen without vibration.

Fig. 3.



The load-band displacement diagram.

2. To measure the band displacement at *one minute* after loading.

3. To remove the load at once. The specimen is left unloaded for *three minutes*.

Repetition of such a procedure gives a system of data showing the relation between w and x.

3. *Calculation of the Constant.*

In this experiment we get $a=7{\cdot}64$ in revolutions of micrometer, $b=2$ cm., and $P=5w$ kg. The equation (3) takes the form

$$w=(K_1/19{\cdot}1)x-P_0/5.$$

Plotting w to x on the rectangular coordinates, we have the load-band displacement diagram shown in fig. 3.

Drawing a straight line through the points plotted, we have

$$m=0{\cdot}67 \quad \text{and} \quad K_1=12{\cdot}8 \text{ kg./cm.}$$

The coefficient of extinction is now determined, the creep of *one minute* being taken into account. For the same data the method of least squares gives

$$K_1=13{\cdot}0 \text{ kg./cm.},$$

which is slightly different from the previous value determined graphically.

V. Beam Method.

1. *Theory A* (*Multiple Stage Method*).

The stress may be given by the formula $S=3yP/d$ at the point y distant from the neutral axis of the standard beam shown in fig. 1. Denoting the stresses corresponding to the loads P_1 and P_2 by S_1 and S_2 respectively, we have

$$S_1-S_2=3y\Delta P/d,$$

where $\Delta P=P_1-P_2$. If N_1 and N_2 are the orders of extinction for S_1 and S_2 respectively, we get

$$S_1-S_2=K_t\Delta N/d,$$

where $\Delta N=N_1-N_2$, the creep of t minutes being taken into account. From these two expressions we have

$$K_t=3y\Delta P/\Delta N. \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Thus the coefficient of extinction is equal to $3y$ multiplied by the ratio of the load-increment to the corresponding increment of the orders of extinction. It characterizes the above procedure that the optical effect at a point is to be measured at least on two stages of the load.

When the initial stress S_0 exists at the point y, it may be supposed that N corresponds to the stress $(S+S_0)$. As S_0 has no effect upon the value of (N_1-N_2), the initial stress adds nothing new to the equation (4).

The constant load due to the lever and its accessories may be neglected in this case, too.

2. *Theory B (Single Stage Method).*

Denoting the stresses at the points y_1 and y_2 distant from the neutral axis of the beam, shown in fig. 1 by S_1' and S_2' respectively, we have

$$S_1'-S_2'=3P\,\Delta y/d,$$

where $\Delta y=y_1-y_2$, the load P being kept constant in this case. If the orders of extinction for these stresses are N_1' and N_2' respectively, we get

$$S_1'-S_2'=K_t\,\Delta N'/d,$$

where $\Delta N'=N_1'-N_2'$. From the above two we get

$$K_t=3P\,\Delta y/\Delta N'. \quad . \quad . \quad . \quad . \quad . \quad (5)$$

The ratio $\Delta y/\Delta N'$ represents the tangent of the slope of a $y-N'$ diagram. Only a photograph taken on a single stage of the load is needed to determine the coefficient of extinction. In appearance this method is exceedingly simple. It has, however, some demerits, as follows :—

1. The load P must be the total exact value actually applied.

2. The initial stress must be *nil.*

3. The optical effects at various points are used in this method, while in the multiple-stage one the optical change at a point is measured. The latter is less unreasonable than the former from the theoretical point of view.

4. To make the result sure it is necessary to photograph on several stages of the load even in the single-stage method. Hence it follows that any difference is scarcely found between the two methods regarding the trouble of experiments. Some hesitation is justified in adopting the single-stage method to determine the photo-elastic constant.

3. *Method of Measurement.*

The specimen is held as shown in fig. 4 (Pl. XV.). The isochromatic lines at the centre of the span are photographed on several stages of the load. As in the

case of the compression method, three operations are carried out on each stage of the load, namely :—

1. To apply the load to the beam.
2. To photograph at *one minute* after loading.
3. To unload and leave for *three minutes* as it is.

Fig. 5 (Pl. XVI.) is the result, which shows the fringes for 14 stages of the load, the compression side being upward. Scales are inserted between the strips of bands. The number described in the lower end of each scale shows the order in which the right-hand strip of that scale is photographed. This order and the corresponding weight at the lever end are related as follows :—

Order	1.	2.	3.	4.	5.	6.	7.
Weight	0·5	0·75	1·0	1·25	1·5	1·75	2·0
Order	8.	9.	10.	11.	12.	13.	14.
Weight	2·25	2·5	2·75	3·0	3·25	3·5	3·75

4. *Calculation of the Constant.*

Owing to the existence of initial stresses, the black band of the zero order does not lie in the true centre of each strip, as seen in fig. 5 (Pl. XVI.). This eccentricity has no influence on the determination of the constant, as will be shown later. The orders for the bands above zero order are -1, -2, -3, etc. The orders for the bands below zero order are 1, 2, 3, etc.

The position of each band is measured by the scale inserted. The lower fibre of the beam photographed is taken as the origin of reading, owing to the fact that the true position of the neutral axis is not yet found. Five graduations of the scale are taken as unit (1·0).

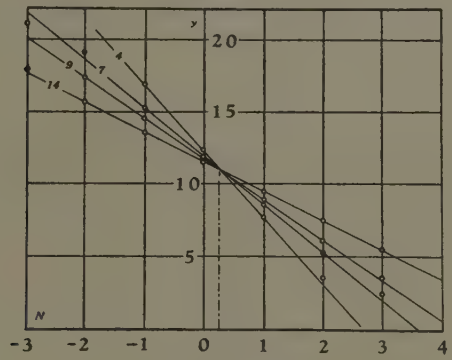
Plotting the scale-reading for the band to the order of extinction, we get a kind of $y-\mathrm{N}$ diagram. Fig. 6 is the one which is drawn for the strips of nos. 4, 7, 9, and 14. The intersection of these lines is the true position of the neutral axis.

The values of N for the lines are measured at the distance 7·0 graduations below the intersection. This distance corresponds to $y=0{\cdot}634$ cm. on the actual beam. The results are as follows :—

Order	4.	7.	9.	14.
Weight w	1·25	2·0	2·5	3·75
N read	1·87	2·44	2·82	3·78

The increments Δw and ΔN are then calculated for several combinations of the lines such as 4—7, 4—9, 4—14, 7—9, 7—14, 9—14. Remembering $P=5w$, the equation (4) gives a value K_1 for each of the above combinations. The mean of them is $K_1=12{\cdot}5$ kg./cm. This is a little smaller than the value determined by the compression method. It is not clear whether the difference is due to the nature of materials or to some defect in measurement. In this respect further investigations are needed.

Fig. 6.



The y—N diagram.

VI. Comparison of the Two Methods.

Merits of the compression method are as follows :—

1. The shape of the specimen is small.

2. Determination of the constant is carried out faster than the beam method.

On the other hand, merits of the beam method are as follows :—

1. No special apparatus, such as Babinet's compensator, is needed.

2. There is little error due to the residual creep, as the measurements by this method are carried out on shorter stages of the load than the compression method.

In a word, each method has its merits and demerits.

VII. Summary.

1. In measuring the optical effect, special attention must be paid to the time at which the optical effect is observed, and the load should be interrupted between two successive observations. These eliminate the error due to *the optical creep*.

2. In calculating the constant from the data experimentally obtained, the increment method should be used. This eliminates the effect of *initial stresses*.

3. In *the compression method* the band displacement x is measured by Babinet's compensator on several stages of the load P. The coefficient of photo-elastic extinction is then given by the relation

$$P=(b/a)\,K_t x-P_0.$$

4. In *the beam method* the isochromatic lines at the centre of the span are photographed on several stages of the load P. The change in the order of extinction N is then examined on the photographs at the point y distant from the neutral axis. The coefficient is determined from the relation

$$K_t=3y\,\Delta P/\Delta N.$$

It is a pleasure to record here a debt of gratitude to the Nippon Gakujutu Sinkokai, which by generous financial support has rendered it possible for the author to carry out the above investigation.

XL. *Entropy, Strain, and the Bose-Einstein Statistics.* *By* W. S. Kimball, *Ph.D.*, *and* Max C. Wygant, *M.S.**

Introduction.

It has already been shown by the new geometrical weight method † that the Maxwell distribution law ‡ and the Fermi-Dirac law § represent equilibrium between stress and strain in velocity space, and thus have their foundation in Newtonian Mechanics and Elasticity Theory. The present paper applies the geo-

* Communicated by the Authors.
† W. S. Kimball, J. Phys. Chem. xxxiii. p. 1158 (1929).
‡ W. S. Kimball, J. Phys. Chem. xxxv. p. 611 (1931).
§ W. S. Kimball and G. Berry, Phil. Mag. ser. 7, xiii. p. 1131 (1932).

metrical weight method to the Bose-Einstein statistics, which is also found to represent equilibrium between stress and strain in velocity space, and thus can be interpreted as resting on a mechanical basis. Not only so, but the geometrical weight method seems to call for a more general statistics which include the Bose-Einstein as a special case.

1. *The Three Stages in the Development of Statistical Mechanics.*

The development of statistical mechanics and the physical concept of entropy may be grouped into three stages. The first postulated that entropy was proportional to the logarithm of the probability, and was developed by Boltzmann. The second, due mainly to Gibbs, measured entropy by extension in phase in connexion with probability. The third stage developed by the Mathematics Department of Michigan State College, disregards probability entirely, and relies solely on Elasticity Theory and Newtonian Mechanics as the basis for Statistical Mechanics. According to this new treatment entropy is proportional to the old-fashioned dimensionless mechanical strain. Thus the first stage is expressed by Boltzmann's equation

$$S = k \log W + \text{const.}, \quad . \quad . \quad . \quad . \quad . \quad (1)$$

which connects the probability W with the entropy S, where k is Boltzmann's constant.

The second stage of development combines the probability or statistical aspect with the *extension in phase* of a system, as used by Gibbs to measure its thermodynamic probability. The method is characterized by a shoving of probability into the background because of its inherent vagueness, and redefining the W of Boltzmann's equation as the " weight " of a system :

$$W = \frac{N!}{n_1!\, n_2! \ldots n_k!} \omega_1^{n_1} \omega_2^{n_2} \ldots \omega_k^{n_k}. \quad . \quad . \quad (2)$$

Eq. (2) represents the culmination of the second period of development. It combines a statistical factor

$$\frac{N!}{n_1!\, n_2! \ldots n_k!}$$

with a geometrical one $\omega_1^{n_1} \omega_2^{n_2} \ldots \omega_k^{n_k}$, the latter

being the product of the cells in phase space, each raised to a power n_i equal to the number of particles it contains. The method of (2) is still statistical, since the n's are taken to be the variable, hence vital part, whereas the phase cells ω_i are kept constant during operations on the weight or probability (2).

The third stage is represented by the new geometrical expression for weight :

$$W = N^N(r_1 r_2 r_3 \ldots r_N). \quad . \quad . \quad . \quad . \quad . \quad (3)$$

This is obtained from (2) by replacing $N!$ by N^N and setting each $n_i = 1$ and each $\omega_i = r_i$. This amounts to nothing more than choosing the cells ω_i to be of such size that each includes a single particle. The cell is then called the *range* of the particle, being the basic physical concept of the geometrical weight method. Thus (3) is physically equivalent to (2). The significance of the step from (2) to (3) lies in the entire removal of *emphasis* from the statistical aspect and placing it on the geometrical factor in the weight of (2) or (3). In fact the constant N^N is all that is left in (3) of the statistical aspect, and the vital, variable part is the geometrical product of the ranges. If we compare the elementary statistical probability W of (1) with the geometrical weight or probability (3), we see that (2) is intermediate in that it involves both statistical and geometrical factors.

Thus a distribution function gives the number of molecules dn in an element $d\omega$ of phase space

$$dn = Nf(c)\,d\omega. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Then the phase volume per molecule or " range " is the reciprocal of the number included in unit volume

$$r_i = \frac{1}{Nf(c_i)} \quad \text{or} \quad W_i = \frac{1}{f(c_i)} = Nr_i, \quad . \quad . \quad (5)$$

where the second expression is the weight per molecule, and the weight (3) becomes

$$W = \frac{1}{f(c_1)f(c_2)\ldots f(c_N)}. \quad . \quad . \quad . \quad . \quad (6)$$

2. *The Geometrical Weight Method transforms Problems of Statistical Mechanics into Problems in Elasticity Theory.*

Notable illustrations of elasticity theory that result from the geometrical weight method are the equilibrium

equations between stress and strain that yield the Maxwell and Fermi-Dirac distributions

$$k\mathrm{T}=\frac{mu_1du_1}{dr_1/r_1}=\frac{mu_2du_2}{dr_2/r_2}=\ldots=\frac{mu_idu_i}{dr_i/r_i},\quad .\quad .\quad (7)$$

which come from maximizing (3) subject to a constant energy E. Hence $k\mathrm{T}$ is the elastic modulus per molecule. If we sum the numerators and denominators of (7), and multiply by n, the number of molecules per c.c., we have for both distributions

$$p=nk\mathrm{T}=\frac{nmu\,du}{dr/r},\quad .\quad .\quad .\quad .\quad .\quad .\quad (8)$$

showing the pressure p as the elastic modulus in velocity space as well as in ordinary space for isothermal changes under the gas law.

From the elasticity point of view it is also striking that entropy is proportional to the strain Y. Thus by (3) and (1)

$$d\mathrm{S}=k\Sigma\frac{dr_i}{r_i},\quad .\quad .\quad .\quad .\quad .\quad .\quad .\quad (9)$$

and

$$\mathrm{S}=k\mathrm{Y}=k\log\mathrm{W}+\text{const.}\quad .\quad .\quad .\quad .\quad (10)$$

Thus Elasticity Theory seems to replace probability and statistics wherever the geometrical weight method finds application in the field allotted to statistical mechanics.

3. *The Bulk Modulus and Stress Strain Relations for Real Gases in Ordinary Space.*

The isothermal bulk modulus is obtained by differentiating the equation of state at constant temperature.

$$p=nk\mathrm{T}=-\frac{dp}{d\mathrm{V}/\mathrm{V}};\quad \text{or}\quad p=nk\mathrm{T}=-\frac{dp}{d\mathrm{V}/(\mathrm{V}-b)}.\quad (11)$$

The first of (11) comes from the gas law and the second from Van der Waals' equation when attractive * forces are neglected. To include these forces take his complete equation and factor out the pressure:

$$p\left(\mathrm{V}+\frac{a}{p\mathrm{V}}-b-\frac{ab}{p\mathrm{V}^2}\right)=\mathrm{RT},\quad .\quad .\quad .\quad (12)$$

* Kimball and Berry, *loc. cit.*

or, neglecting * the second order term ab/pV^2:

$$p\left(V+\frac{a}{pV}-b\right)=p(V+c)=RT, \quad . \quad . \quad . \quad (13)$$

where $c=\dfrac{a}{pV}-b$ is apparently more nearly constant experimentally than either a, b, or $\dfrac{a}{pV}$ considered separately. Thus the table shows a comparison of the relative magnitudes and constancy of these "constants" given by experimental data † for argon gas obtained by the constant volume method. Note how $c=\dfrac{a}{pV}-b$ is

Gas.	T (A).	V (lit.).	R.	p (at.).	a.	$\frac{a}{pV}$.	b.	$\frac{ab}{pV^2}$.	$c=\frac{a}{pV}-b$.
Argon.	151·2	22·4	82·07	1	1900000	84·8	61·0	0·2300	23·8
,,	157·2	22·4	82·07	1	1870000	83·5	59·5	0·2200	24·0
,,	163·2	22·4	82·07	1	1840000	82·2	58·0	0·2100	24·2
,,	173·2	22·4	82·07	1	1785000	79·7	55·5	0·2000	24·2
,,	183·2	22·4	82·07	1	1735000	77·4	53·0	0·1800	24·4
,,	193·2	22·4	82·07	1	1690000	75·4	51·0	0·1700	24·4
,,	213·2	22·4	82·07	1	1600000	71·4	48·0	0·1500	23·4
,,	233·2	22·4	82·07	1	1530000	68·3	45·0	0·1400	23·3
,,	253·2	22·4	82·07	1	1470000	65·6	43·0	0·1300	22·6
,,	273·2	22·4	82·07	1	1420000	63·4	41·0	0·1160	22·4

much more nearly constant than the others over a wide range of temperature.

The isothermal elastic modulus for (13) is obtained as in case (11):

$$p=nkT=-\frac{dp}{dV/(V+c)}=\frac{dF}{F}\Big/\frac{dp}{\left(1+\frac{F}{C}\right)}, \quad . \quad (14)$$

where in the right member we have set

$$V=\frac{N}{F}; \quad \text{and} \quad c=\frac{N}{C} \quad . \quad . \quad . \quad . \quad (15)$$

for the sake of later comparisons.

† L. G. Loeb, 'Kinetic Theory of Gases,' p. 125.
* M. Saha and B. Srivastava, 'Text-book of Heat,' p. 166.

We may compare the strain in (14) with the corresponding entropy change. Take the first law of thermodynamics:

$$dQ = dU + pdV. \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

Let $dU = 0$ for isothermal changes, and divide by T and introduce (13):

$$dS = \frac{dQ}{T} = \frac{pdV}{T} = R\frac{dV}{V+c}. \quad . \quad . \quad . \quad (17)$$

Eq. (17) shows that these isothermal entropy changes are proportional to the strain involved in (14).

4. *Entropy, Strain, and Weight for Real Gases in Ordinary Space.*

For the range in ordinary space we depart from $1/n = 1/F$ applicable* where the gas law applies, and also from

$$\frac{1}{n} - b_1 = \frac{1}{F} - \frac{1}{B},$$

which restricts the range per molecule by the effective volume it displaces †. Instead we *increase* its range, using (15):

$$r = \frac{1}{n} + c_1 = \frac{1}{F} + \frac{c}{N} = \frac{1}{F} + \frac{1}{C}. \quad . \quad . \quad . \quad (18)$$

Thus for practical and mechanical purposes the attractive Van der Waals's forces are taken to amount to an increase c_1 in the volume $1/n$ allotted to each molecule, in much the same way as the decrease in volume by $b_1 = \frac{b}{N}$ amounts to a repulsive force, arising from extra crowding. The weight per molecule then becomes

$$w = Nr = \frac{N}{n} + Nc_1 = V + c = \frac{N}{F}\left(1 + \frac{F}{C}\right), \quad . \quad (19)$$

and for N molecules

$$W = (V+c)^N = \left(\frac{N}{F}\right)^N \left(1 + \frac{F}{C}\right)^N. \quad . \quad . \quad . \quad (20)$$

* Kimball, *loc. cit.*
† Kimball and Berry, *loc. cit.*

Hence for the entropy

$$\left.\begin{aligned} S&=Nk\log(V+c)+\text{const.},\\ dS&=R\frac{dV}{V+c}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (21)$$

Comparison of (14) and (17) with (18) to (21) shows how entropy and strain are related to range and weight in ordinary space.

5. *Entropy, Strain, and Weight for Real Gases in Velocity and Momentum Space.*

For real gases the number of molecules in unit volume of velocity or momentum space is represented as in the classical statistics by a distribution function

$$\frac{dN}{dw}=f(u,v,w)=n_i=\frac{1}{V_i}. \quad . \quad . \quad . \quad . \quad . \quad (22)$$

Here, however, we do not take its reciprocal for the range as in the classical statistics, but for neutral or attracting molecules this reciprocal is to be *increased*, as done for ordinary space by (18), by an amount d_1 per molecule, in contrast with the Fermi-Dirac statistics and the Pauli exclusion principle *, which involve a *decrease* in range;

$$r_i=\frac{1}{n_i}+d_1=\frac{1}{f}+\frac{1}{D}=\frac{1}{f}\left(1+\frac{f}{D}\right)=\frac{1}{g_i}. \quad . \quad (23)$$

The weight per molecule is then

$$w_i=Nr_i=\frac{N}{f}\left(1+\frac{f}{D}\right)=\frac{N}{g_i}, \quad . \quad . \quad . \quad . \quad (24)$$

and for N molecules the weight is

$$W=N^N(r_1r_2\ldots r_N)=N^N\prod_{i=1}^{N}\left(\frac{1+\frac{f_i}{D}}{f_i}\right)=\frac{N^N}{g_1g_2\ldots g_N}, \quad (25)$$

and the entropy appears again as k times the sum of the strains:

$$S=k\sum_1^N\log r_i+C=k\sum_1^N\int\frac{dr_i}{r_i}=kY. \quad . \quad . \quad (26)$$

Eq. (24) and (25) differ from the corresponding expressions given by Chandrasekhar † only in that the constant

* Kimball and Berry, *loc. cit.* eq. (20).

† S. Chandrasekhar, Phil. Mag. lx. p. 621 (1930), eq. (8).

D has replaced his A (number of cells per unit volume in phase space) that appears in the Bose-Einstein statistics.

6. *The Geometrical Weight Method gives the Bose-Einstein Statistics as a Special Case.*

To obtain the velocity distribution we apply the Lagrange method, and maximize (25) subject to a constant energy E by setting equal to zero the partial derivatives of $\phi = W + \lambda E$ with respect to the velocity components, where λ is an undetermined constant:

$$\frac{\partial \phi}{\partial u_i} = \frac{\partial w}{\partial u_i} + \lambda \frac{\partial E}{\partial u_i} = -W \frac{d}{du_i} \log g_i + \lambda m u_i = 0, \quad (27)$$

or by (23):

$$W \frac{dr_i}{r_i} + \lambda m u_i du_i = 0, \quad . \quad . \quad . \quad . \quad (28)$$

or

$$-\frac{W}{\lambda} = kT = \frac{m u_1 du_1}{dr_1/r_1} = \frac{m u_2 du_2}{dr_2/r_2} = \cdots$$

$$= \frac{m u_i du_i}{dr_i/r_i} = -\frac{m u\, du}{\frac{df}{f}\left(1 + \frac{f}{D}\right)} \cdot \quad . \quad . \quad . \quad (29)$$

Eq. (29) show how the critical eq. (27) and (28) involve a constant ratio between energy differences per molecule and the corresponding strain, whereby it appears that $-\frac{W}{\lambda}$ (required * by thermodynamics to be kT) is the elastic modulus per molecule for the indicated stress-strain relations. If we sum the numerators and denominators and multiply both members by n we get

$$p = nkT = n \frac{\Sigma m u_i du_i}{\Sigma \frac{dr_i}{r_i}} = \frac{n m u\, du}{dr/r} = -\frac{n m u\, du}{\frac{df}{f}\left(1 + \frac{f}{D}\right)} \cdot \quad (30)$$

The numerator of (30) is the energy difference per unit volume, so that the pressure p again appears as the elastic modulus connecting stress and strain in velocity space, just as it is, by (14), for real gases in ordinary space.

* J. Rice, 'Statistical Mechanics,' Appendix on Recent Developments.

The integration of (29) gives

$$f = \frac{1}{\mathrm{C}e^{\frac{mu^2}{2k\mathrm{T}} - \frac{1}{\mathrm{D}}}}, \quad \ldots\ldots \quad (31)$$

where C is the constant of integration. This is a distribution more general than the Bose-Einstein statistics by virtue of the general character of the constant D. If we set this D = A, the number of cells per unit volume in phase space, (31) becomes the Bose-Einstein statistics :

$$f = \frac{1}{\mathrm{C}e^{\frac{mu^2}{2k\mathrm{T}} - \frac{1}{\mathrm{A}}}}. \quad \ldots\ldots \quad (32)$$

7. *The Two Points of Similarity between the Behaviour of Gas in Ordinary Space and Velocity Space*, i. e., *Elastic Modulus and Range.*

For the case of neutral or attracting molecules we may emphasize in this new domain, as already done for ideal gas * and repelling molecules †, how there is a striking parallelism of molecular behaviour in ordinary and in velocity space.

As to elastic moduli, (14) parallels (29) and (30), since each gives the simple pressure formula $p = nk\mathrm{T}$ as the elastic modulus connecting stress (energy difference per unit volume) and strain, whereas $k\mathrm{T}$ may be thought of as the elastic modulus per molecule.

As to range, eq. (18) parallels (23), each being the reciprocal of the number of molecules per unit volume plus (since we treat neutral and attracting molecules) an undetermined constant written as c_1 or d_1 respectively.

Now range is very intimately connected with strain and entropy by (18) to (21) and (23) to (26), which are mechanical quantities functioning about an elastic modulus. Hence this parallelism of behaviour in ordinary and velocity space suggests the possibility of its having some underlying mechanical basis, as though there were a mechanical reason why molecules should "object" to being crowded together with the same vim whether in velocity or ordinary space.

* Kimball, *loc. cit.*
† Kimball and Berry, *loc. cit.*

8. *The Difference between the Applications of the Geometrical Weight Method to the Bose-Einstein and the Fermi-Dirac Statistics.*

In using the geometrical weight method to derive the Fermi-Dirac statistics the Pauli exclusion principle was invoked * to reduce the range of a molecule by the constant amount of a single cell in phase space, being the minimum range it could occupy. This represented an exact parallelism of method in phase and ordinary space, since the constant b of Van der Waals's equation is justified by the same line of reasoning. Thus the similarity of form in the case of the *reduced* ranges presents a real physical similarity, in that the constants A and B each represents the number of cells per unit volume of the space referred to.

In case of the Bose-Einstein statistics, however, there is a contrast rather than parallelism in respect to the constants A of (32) and C of (18) to (20), in that A, the number of cells per unit volume, implies an increase in range per molecule by an amount $a_1 = \frac{1}{\mathrm{A}}$ equal to a single cell, whereas the C of (18) to (20) implies an increase per molecule equal to $c_1 = \frac{c}{\mathrm{N}} = \frac{1}{\mathrm{C}}$, where c is the experimental $\frac{a}{p\mathrm{V}} - b$ of Van der Walls's equation (13) instead of his constant b. To perfect the parallelism for the Bose-Einstein as is done for the Fermi-Dirac statistics the above c of (13) would have to be replaced by Van der Waals's b, which is not justified by experiment.

9. *Parallelism of Behaviour in Ordinary Space and Phase Space as a Guide to Generalization.*

Heretofore the parallelism in mechanical behaviour of physical quantities like elastic moduli, range, strain, etc. in ordinary and in velocity space has been universal and complete. Hence we may let it be a guide to our thought, and call it tentatively a physical *principle* that results from a closer identity between actual conditions in ordinary and phase space than appears on the

* Kimball and Berry, *loc. cit.* eqs. (20) and (24).

surface. Ehrenfest has already considered * and dismissed † the idea of a *causal* relationship between Van der Waals's exclusion constant and the Pauli exclusion principle, but nevertheless a general parallelism seems significant.

Various generalizations ‡ of the new statistics have been suggested, and if the above parallelism is significant it might be a guide to their validity. This latter should accordingly depend on how they fit into the elastic relations of the geometrical weight method and to what extent these relations in phase space parallel the corresponding relations in ordinary space.

Following this thought the principles and ideas of the geometrical weight method, elasticity theory, and Newtonian mechanics as applied in this paper call for a more general distribution (31), whose constant D corresponds to the more general experimental c of (13), but may be taken to include the A of the Bose-Einstein statistics as a special case. Thus, viewed as a distribution function (31) fits experiment (at the present writing) as well as (32), and according to the above is more in accord with mechanical principles.

Michigan State College,
East Lansing, Michigan.

XLI. *The Application of a modified Morse Formula to simple Hydride Diatomic Molecules (Di-atoms).* *By* C. H. Douglas Clark, *M.Sc., A.R.C.S., D.I.C., Assistant Lecturer in Chemistry in the University of Leeds* §.

I. *Introductory Note on Nomenclature : Suggested Use of the Term "Di-atom."*

IN accordance with modern views concerning the electronic structure of molecular groups where two atoms only are present, three types may be recognized, as set forth in the scheme below [1], [2].

* P. Ehrenfest, 'Nature, cxix. p. 196 (1927).
† *Loc. cit.* p. 602 (1927).
‡ R. Lindsay, "A Modification of Brillouin's Unified Statistics." Phil. Mag. xvii. p. 264 (1934); L. Brillouin, 'Quantum Statistics.'
§ Communicated by the Author.

Class.	Name.	Type.	Example.
I *a*. ..	Polar.	Separated-ions.	NaCl.
I *b*. ..	Non-hydride homopolar.	Separated-atoms.	NO.
II. ..	Hydride.	United-atom.	FH.

Each of the three classes has its own peculiar chemical and spectroscopic characteristics. It appears that a new term may, at the present stage of investigation, conveniently be suggested to cover these different cases. This point of view appears strengthened by the following considerations :—

(1) The term " diatomic molecule," when frequently repeated, especially when associated with the qualifications " hydride " or " non-hydride," tends to become cumbersome.

(2) The atomic groupings conventionally designated by the term are generally, though not invariably, unknown as stable chemical molecules, and appear to defy the ordinary rules of valency : *e. g.* BeF, BeO, SiN, etc. Herzberg [3] prefers to say that such groups possess physical rather than chemical stability. Further, the term " diatomic molecule " does not seem very appropriate to describe cases of united-atoms (Class II. above). Some new term appears necessary to include homopolar hydrides, such as CH, BH, etc., where the grouping can hardly be described as truly molecular.

(3) As the study of polyatomic molecules proceeds, it becomes increasingly realized that the spectroscopic properties of molecular groups where two atoms only are present are entirely characteristic of the simplest case where one atomic linkage only is present.

Langmuir [4] used the term " co-molecule " to represent the general case of any group of atoms held together by electron pairs, whether the group was charged or uncharged. This term is unsuitable in the present connexion, however, since it does not particularly specify diatomic groups. Amongst other possibilities, " dipole " implies electrical asymmetry, which may not be present, whilst " doublet " is used in describing spectral mul-

tiplicity. The word "diatom," by a somewhat unfortunate coincidence, is already appropriated in biological science, although it is quite unlikely to be confused in practice with use in the present connexion. Further, "biatom" is, perhaps, not very attractive, and is etymologically hybrid.

After careful consideration of the position, and discussion with his colleague Mr. R. S. Bradley, the author would propose the term "di-atom" (with hyphen), an ionized group being described by "di-atom ion." This nomenclature is adopted throughout the present paper.

It has been further considered convenient to use the symbol FF in place of the conventional F_2, and similarly in like cases.

II. *The Modified Formula for simple Hydride Di-atoms.*

It is found that the modified Morse expression [5]

$$\omega_e r_e^2 . \sqrt{n} = k - k' = \mathrm{K}, \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where n is the group number of the di-atom, or number of shared electrons, and K is a "period constant,"

TABLE I.

Period Constants in modified Morse Expression for Hydride Di-atoms.

Period.	K($\times 10^{-5}$).	Period.	K (singly ionized, positive ions) ($\times 10^{-5}$).
KH	8·05	KH^+	(7·05)
LH.....	15·25	LH^+	13·20
MH	(18·8)	MH^+	....

can be applied directly to neutral hydride di-atoms (Class II.) of the KH period from LiH to FH with generally satisfactory agreement with experiment. Data for only one di-atom ion (BeH^+) of this period are recorded by Jevons [12], so that the value of K is uncertain. It then becomes possible to assign the period constants for hydrides shown in Table II. (Only one di-atom of

the MH period (CaH) has the necessary data recorded, so the period constant is again doubtful.)

III. *Hydride Di-atoms of the KH Period.*

The modified Morse expression for neutral hydride di-atoms takes the form derived from equation (1) and the period constant from Table I., as follows :—

$$\omega_e r_e^3 . \sqrt{n} = 8{,}050, \quad . \quad . \quad . \quad . \quad . \quad (2)$$

TABLE II.

Equilibrium Internuclear Distances in Neutral Di-atoms of the KH Period.

Hydride di-atom.	Periodic group. n.	Electronic state.	ω_e (cm.$^{-1}$).	r_e (Å.U.). (expt.).	r_e (Å.U.). Morse (calc.).	r_e (Å.U.). Clark (calc.).	Percentage errors. (M).	Percentage errors. (C).
LiH	2	$G\ ^1\Sigma^+$	1406·1	1·6	1·622	1·593	+ 1·3	−0·4
BeH	3	$^2\ \Pi_{reg.}$	2087·6	1·330	1·421	1·305	+ 6·8	−1·9
		$G^2\Sigma^+$	2058·5	1·340	1·429	1·312	+ 6·6	−2·1
BH	4	$^1\Pi$	(2450)	1·213	1·348	1·180	+ 11·1	−2·7
		$G\ ^1\Sigma^+$	(2230)	1·226	1·391	1·217	+13·5	−0·7
OH	7	$^2\Sigma^+$	3182·5	1·009	1·235	0·986	+20·7	−2·1
FH	8	$G\ ^1\Sigma$	4037	0·864	1·141	0·890	+32·1	+3·0
Average percentage errors (irrespective of sign)							13·2	1·8

G.—Ground state.

where r_e is in Å.U. Calculation is compared with experiment for hydrides from LiH to FH, so far as data are available, in Table II.

In Table II. the errors using Morse's original expression become increasingly positive with increasing group number, as noticed earlier in the case of non-hydride di-atoms. The modified relation gives much better agreement, and the errors do not progress steadily in one direction. (For one excited state of LiH the data are not included. The error using Morse's relation is about

10 per cent., which does not fit in with the gradation in errors noticed above. Hence, probably, the estimate of r_e is at fault; it is given in Jevons's table to one decimal place only.) The rather large error in the excited level of BH may be due to uncertainty in ω_e. The appearance of a positive 3 per cent. error in the case of FH is striking. The difference error, however, is 0·026 Å.U., of the same order of magnitude as obtained in other cases. The average error is reduced from 13·2 per cent. to 1·8 per cent.

TABLE III.

Percentage Diminutions of Internuclear Distance in the KH Period.

Hydride di-atom XH.	r_e(XH) ground state (Å.U.).	r(X) (Å.U.).	r(X)+r(H) (Å.U.).	$\frac{100r_e(\mathrm{XH})}{[r(\mathrm{X})+r(\mathrm{H})]}$.
LiH	1·59	1·50	2·03	78·3
BeH.....	1·34	1·14	1·67	80·2
BH.....	1·23	0·97	1·50	82·0
OH.....	1·01	0·70	1·23	82·1
FH.....	0·86	0·68	1·21	71·1

It is convenient to define the "percentage diminution of internuclear distance" by the expression

$$100r_e(\mathrm{XH})/[r(\mathrm{X})+r(\mathrm{H})],$$

where r(X), r(H) are the radii of the constituent atoms. This function is found to reach a maximum near the middle of the KH period (see Table III. above). The atomic radii are taken from a useful table due to Sidgwick [6], except the estimate of the radius of the boron atom, which is due to Hägg [7]. The initial increase of percentage diminution may be associated with increasing repulsion between the X and H nuclei as the charge of X increases from Li to B. The subsequent diminution of the percentage function may, perhaps, be associated with increasing shielding of the nuclear charge of the atom X. The value of r(H) is taken as 0·53 Å.U.

IV. *Neutral and Ionized Di-atoms of the LH Period.*

When the modified Morse expression is applied directly to determination of r_e values of MgH, AlH, ClH, MgH^+, and ClH^+, the average percentage errors are reduced from 6·2 per cent. (Morse) to 3·9 per cent. (Clark). Using the period constants from Table II., the equation in the case of neutral di-atoms of the LH period is

$$\omega_e r_e^3 . \sqrt{n} = 15{,}250. \quad . \quad . \quad . \quad . \quad . \quad (3)$$

and for ionized di-atoms

$$\omega_e r_e^3 . \sqrt{n} = 13{,}200. \quad . \quad . \quad . \quad . \quad . \quad (4)$$

It must be admitted that the agreement between theory and experiment is not as good in the LH period as in periods previously examined. The influence of deformation is difficult to allow for, and there is no available rule by which one may decide in these cases exactly how many electrons should be regarded as shared. Hence the appropriate value of n is somewhat uncertain. Empirically it seems necessary to allow for a number of shared electrons in earlier groups greater, and a number in later groups smaller, than is derived from adding together the valency electrons of the atoms concerned. This arises since the errors using n become increasingly negative with increasing group number in both the periods of neutral and ionized di-atoms. An "effective atomic number" n' may be introduced, derived from n by the following rule:—"For neutral hydride di-atoms of the LH period, when $n=4$, no change is made: $(n'=n)$; if $n<4$, n' is the next EVEN number above n: $(n'>n)$; if $n>4$, n' is the next EVEN number below n: $(n'<n)$. For di-atom ions of the period, substitute ODD for EVEN." The results are shown in Table IV., using both n and n'.

With the single exception of an excited state of MgH^+, the errors using the effective group number are reduced as compared with those using the regular group number. It seems impossible to say at the present stage, with the available data, what significance (if any) should be attached to the use of an even effective group number for neutral hydrides and of an odd effective group number for ionized hydrides in this period. It can only be placed on record that the modification fits the facts,

TABLE IV.

Equilibrium Internuclear Distances in Neutral Di-atoms and Di-atom Ions of the LH Period.

Neutral hydride di-atom	Hydride di-atom ion.	Periodic group numbers.		Electronic state.	ω_e. (cm.$^{-1}$).	r_e(Å.U.).				Percentage errors.		
						(expt.)	Morse (calc.).	Clark (calc.).				
		n.	n'.					n.	n'.	M.	C(n).	C(n').
MgH	—	3	4	B ^{2}	(1700)	**1·67**	1·523	1·730	**1·649**	−8·8	+3·6	**−1·3**
				A $^2\Pi_{reg.}$	1603·5	**1·68**	1·552	1·764	**1·682**	7·6	+5·0	+0·1
				G X $^2\Sigma^+$	1493·5	**1·73**	1·590	1·806	**1·722**	−8·1	+4·4	**−0·5**
AlH	—	4	4	G A $^1\Sigma^+$	1680·6	**1·644**	1·529	1·655	**1·655**	−7·0	+0·7	**+0·7**
ClH	—	8	6	G $^1\Sigma$	2989·7	**1·272**	1·261	1·217	**1·285**	−0·9	−4·3	**+1·8**
—	MgH^+	2	3	$^1\Sigma^+$	1138·4	**2·008**	1·740	2·017	**1·885**	−13·4	+0·5	**−6·2**
				G $^1\Sigma^+$	1702·2	**1·649**	1·523	1·764	**1·648**	−7·6	+7·0	**−0·1**
—	ClH^+	7	5	$^2\Sigma^+$	1603	**1·528**	1·552	1·460	**1·544**	+1·6	−4·4	**−0·5**
				G $^2\Pi_{inv.}$	(2608)	**1·313**	1·320	1·242	**1·313**	+0·5	−5·3	**0·0**
Average percentage errors (irrespective of sign)												
							Neutral di-atoms			6·5	3·6	**0·9**
							Di-atom ions			5·8	4·3	**1·7**
							Total average errors			6·2	3·9	**1·2**

G.—Ground state.

the average Morse error of 6·2 per cent. being reduced to 1·2 per cent. in the process.

It does not appear possible definitely to decide, with the available data, whether a maximum percentage diminution of internuclear distance occurs in this period.

V. *Morse's Function and Electron Polarization.*

Debye [8] found that, where two atoms having charges $+e$ and $-e$ and polarizabilities α_1 and α_2 respectively are located at a distance r_e apart, the dipole moment is given by

$$\mu = er_e\left[\frac{1-\frac{\alpha_1}{r_e^3}+\frac{\alpha_2}{r_e^3}+\frac{4\alpha_1\alpha_2}{r_e^6}}{1-\frac{4\alpha_1\alpha_2}{r_e^6}}\right], \quad . \quad . \quad . \quad (5)$$

where e is the electronic charge. In the case of a hydride XH, the polarizability of the hydrogen ion is zero, so that $\alpha_1=0$ and α_2 may be set equal to α, the polarizability of the X ion. Equation (5) then reduces to the simpler form

$$\mu = er_e\left[1-\frac{\alpha}{r_e^3}\right] = \mu'\left[1-\frac{\alpha}{r_e^3}\right], \quad . \quad . \quad . \quad (6)$$

where $\mu'=er_e$. The electron polarization P may then be obtained by means of the Debye expression

$$P=\frac{4}{3}\pi N\alpha. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Further, since $\mu=et$, where t is the dipole distance, equation (6) may be written in the form

$$\frac{t}{r_e}=1-\frac{\alpha}{r_e^3} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

or

$$\alpha = r_e^3\left[\frac{r_e-t}{r_e}\right] = r_e^3\left[\frac{\mu'-\mu}{\mu'}\right]. \quad . \quad . \quad . \quad (8\,a)$$

Debye showed, by an application of the numerical data, that α is really an "apparent polarizability," in the cases of ClH and BrH, very approximately yielding polarizations P equal to the atomic refractions of argon and krypton respectively, the inert gases having the

same number of extranuclear electrons as the two halides. This seems to indicate that r_e^3 is a measure of this apparent polarizability in terms of what may be called "the relative difference between internuclear and dipole distance" (compare equation $(8\,a)$ above). It is well known that electron polarization P is connected with "true atomic (or molecular) volume," and has the dimensions of the cube of length.

It may, therefore, be suggested that r_e^3 enters into Morse's function for hydrides, and possibly in other cases also, on account of polarization, and that the function represents a magnitude proportional to equilibrium nuclear frequency and apparent polarizability conjointly. This may be found to account for the observed periodicity of Morse's function, studied in another communication.

Assuming that the apparent polarizability of F in FH is equal to the atomic refraction of neon, by Debye's method, the present author has formed an estimate of the dipole moment of the unassociated gas, a quantity which would be extremely difficult to measure directly [9]. The calculation gives $\mu = 1{\cdot}58$ Debye units, a figure which fits in quite well with the known tendency of the gas to associate, and with its position in the periodic groups. These results will be communicated elsewhere in due course.

An observation of v. Wiśniewski [10], [11] to the effect that

$$\sqrt{\epsilon} - 1 = kr_e^3, \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where ϵ is the dielectric constant of a gaseous di-atom and k is a proportionality constant, again suggests a relation between the cube of internuclear distance and dielectric or polarization phenomena. The applicability of this relation will be carefully examined.

VI. *Summary.*

It is suggested that the term "di-atom" may conveniently replace the conventional "diatomic molecule" in spectroscopic nomenclature.

The modification of Morse's formula proposed by the author is found to work satisfactorily for hydrides of the KH period and, with certain restrictions, for hydrides of the LH period also.

The further suggestion is made that the cube of internuclear distance enters into Morse's relation on account of polarization phenomena.

References.

(1) R. S. Mulliken, 'Chemical Reviews,' ix. p. 348 (1931).
(2) C. H. Douglas Clark, 'Chemical Reviews,' xi. p. 231 (1932).
(3) G. Herzberg, 'Leipzige Vortäge,' p. 167 (1931).
(4) I. Langmuir, Journ. Amer. Chem. Soc. xli. p. 868 (1919).
(5) C. H. Douglas Clark, Phil. Mag. (vii.) xviii. p. 459 (1934).
(6) N. V. Sidgwick, 'The Covalent Link in Chemistry' (Cornell University Press, 1933). (See table xviii. on p. 85.)
(7) G. Hägg, *Zeit. physikal. Chem.* B. xi. p. 152 (1930).
(8) P. Debye, 'Polar Molecules' (Chemical Catalog Co., 1929). (See p. 59.)
(9) C. H. Douglas Clark, *Physikal. Zeit.* xxxv. p. 622 (1934).
(10) F. J. v. Wiśniewski, *Zeit. Physik*, xxxvii. p. 126 (1926).
(11) F. J. v. Wiśniewski, *Zeit. Physik*, xliv. p. 392 (1927).
(12) W. Jevons, 'Report on Band-Spectra of Diatomic Molecules' (Cambridge University Press, 1932). (See Appendix II. on p. 280.)

Department of Inorganic Chemistry,
The University of Leeds.
November 8th, 1934.

XLII. *One-dimensional Electron Diffraction.*

By J. R. Tillman, *Beit Scientific Research Fellow* *.

[Plates XVII. & XVIII.]

THE possibility of one-dimensional diffraction of fast electrons by single crystals has been demonstrated by Emslie † in the case of stibnite. With a particular setting of the (010) face of this crystal the plane of incidence can be made to contain a row of antimony atoms, regularly spaced 11·4 Å. apart. The diffraction pattern when photographed contains circles concentric with the point of intersection of this row of atoms and the photographic plate. Emslie's explanation of the circles was as follows :—

"But, to obtain the one-dimensional diffraction that the circles seem to be, an electron must have its wave packet confined chiefly to a single row of atoms parallel to the *a*-axis. The potential in such a region may be high.

* Communicated by Prof. G. P. Thomson, M.A., F.R.S.
† Phys. Rev. xlv. p. 43 (1934).

"The mechanism whereby an electron wave packet becomes localized may be as follows. The original plane wave electron loses a little energy to an atom and becomes an approximately spherical wave radiating from that atom. If it leaves the atom in a direction nearly parallel to the a-axis it may continue in that direction, entrapped in a kind of potential tube. It is scattered finally by the potential variations along the length of the tube."

He supported this statement by measurements made on the radii of the circles, and showed that the mean potential existing inside this tube was about 26 volts higher than that outside the crystal. We will return to Emslie's explanation later, after the results of the present investigation have been fully presented.

The chief reason why such circles had not previously been observed lies in the smallness of the spacings associated with those crystals usually employed in electron diffraction work, such as rock-salt and galena. The circles were therefore too large to be included in the field of view. However, the present paper deals with another crystal—zincblende—frequently used in this kind of work, which, when cleaved along the (110) face, can be made to give moderately large spacings between its zinc atoms in certain directions in this face. Four separate circles have been obtained; from measurements of their radii, together with a knowledge of the wave-length of the primary electrons and the spacings of the atoms, values of the mean potentials existing inside certain tubes have been calculated. These tubes extended in different directions in the crystal surface and contained both zinc and sulphur atoms. The separate values have been compared and some estimates made of them from considerations of the electrostatic fields existing around the zinc and sulphur nuclei in the crystal.

Theory.

We will now consider the process of one-dimensional diffraction by atoms regularly spaced a distance d apart, their direction being perpendicular to the photographic plate. The electron must be assumed to be moving along this row of atoms inside a tube whose mean value of electrostatic potential is Φ. Fig. 1 shows the directions of the primary, secondary or incident, and diffracted

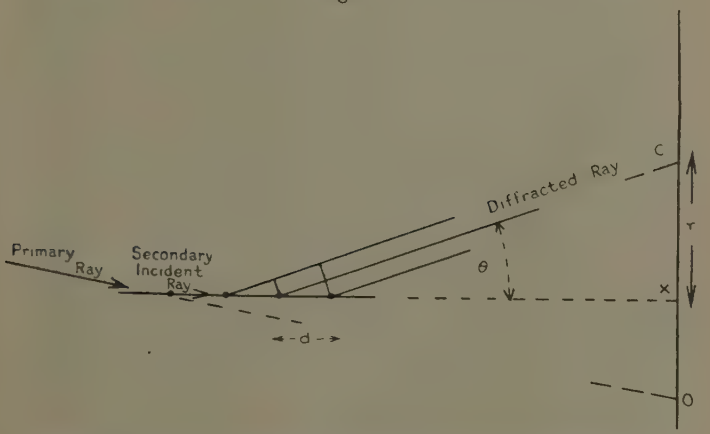
rays, the row of atoms and the photographic plate. The primary ray, which leaves its mark on the plate at O, is of little importance here, except in so far as it gives rise to the incident ray.

The condition necessary for a diffraction maximum is

$$\frac{2\pi d}{\lambda'} - \frac{2\pi d \cos\theta}{\lambda} = 2\pi, 4\pi, \text{ etc.},$$

where λ represents the wave-length of the electron *in vacuo* and λ' in the potential tube. Considering only the first order, we see that this condition implies a circle with centre at X and of radius XC. Since $\lambda = \lambda'\sqrt{1+\Phi/P}$,

Fig. 1.



where P is the voltage associated with the incident ray, the above equation reduces to

$$\frac{\Phi}{P} = \frac{2\lambda}{d}\left(\cos\theta + \frac{\lambda}{2d}\right) - \sin^2\theta. \quad . \quad . \quad . \quad (1)$$

This relation enables Φ to be calculated, λ, d, P, and θ being known. Neglecting for the moment Φ/P, which is small, we should expect

$$\sin^2\theta = \frac{2\lambda}{d}\left(\cos\theta + \frac{\lambda}{2d}\right).$$

This equation differs slightly from that quoted in Emslie's paper, where $\lambda/2d$ was ignored in comparison with $\cos\theta$

(which was taken as unity) and tan θ was taken as equal to sin θ.

Experimental.

The apparatus used has already been described in a previous paper *; a smaller anode hole, 0·15 mm. in diameter, was fitted. Equation (1) shows the necessity for accurate values of λ. For this purpose the method of measuring P by means of a spark-gap and then calculating λ was discarded; but it can be shown that if i be the current flowing through the coil acting as a magnetic lens in the apparatus * when the beam of electrons is sharply focussed then λi = a constant. Specimens of known lattice dimensions were used to determine this constant, separate values of which were consistent to within $\pm 0{\cdot}6$ per cent. The mean value of this constant was then used to give subsequent values of λ. But to obtain more accurate values of the wave-length the diffraction photographs themselves were used. Emslie † has shown how measurements of radii of circles of spots and of distances between rows of spots can give relations between λ and the lattice dimensions. These methods have been used here, assuming the dimensions of the unit cell of zincblende. Unless the separate determinations of λ for any one plate have agreed to within less than 0·3 per cent. that plate has been discarded. Most of the measurements on the plates have been made with a low-powered travelling microscope. In this way distances could be measured accurately to within ·02 mm.

The P-patterns accompanying the semicircles are of some importance. They proved useful in establishing the azimuth. The widths of the bands and their directions with respect to the central band were measured, and the spacings of the planes to which they corresponded calculated. In this way no doubt was left with regard to the azimuth in use.

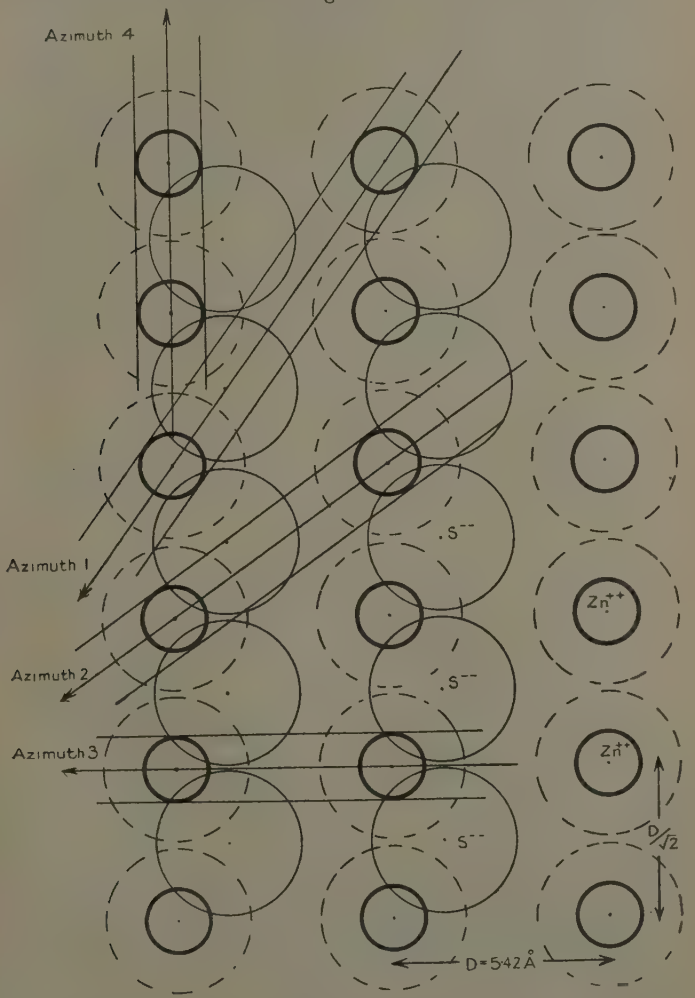
Results.

It is necessary in the first place to say a little about the structure of zincblende. This crystal belongs to the

* Phil. Mag. xviii. p. 656 (1934).
† *Loc. cit.*

cubic system. The zinc atoms lie on a face-centred lattice, with the sulphur atoms on a lattice exactly similar

Fig. 2.



in structure and size but with its origin at the point $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ in the zinc lattice. The side, D, of the unit cell

is 5·42 Å. This naturally occurring crystal cleaves along the (110) face, the distance between successive cleavage planes being $1{\cdot}91_5$ Å. The relative positions of the zinc and sulphur atoms in the surface of this face are important for our purpose, and are shown in fig. 2. We are not, however, concerned with the lack of symmetry of the zinc and sulphur atoms except in so far as it concerns the alignment of atoms in oblique directions across the surface. Since the sizes of the two ions, Zn^{++} and S^{--}, are of importance later on, these have been indicated in the figure, the zinc by the small complete circles and the sulphur by the large. All the atoms depicted lie in the same plane: were the atoms in the plane immediately below this one included in the diagram they would show an exactly similar structure, moved, however, a distance D/2 to the right or left. The four azimuths, which we now consider, are also shown.

TABLE I. (*a*).

Spacing calculated in Å.	Angle measured.	Indices assigned.	Calculated angle.	Spacing in Å, assuming D = 5·42 Å.
1·09	0°	242	0°	$1{\cdot}10_5$
1·88	30°	220	30°	$1{\cdot}91_5$
$0{\cdot}94_5$	30°	440	30°	0·96
$1{\cdot}07_5$	60°	$42\bar{2}$	60°	$1{\cdot}10_5$

TABLE I. (*b*).

λ in Å.	Φ_1 (volts).	λ in Å.	Φ_1 (volts).
·0970	15·5	·0716	10·0
·0893	20·0	·0704	21·0
·0824	16·5	·0703	22·0
·0808	16·5	·0703	21·0
·0740	22·5	·0685	16·0
·0732	19·5	·0682	18·0
·0732	16·0		
		Mean value of Φ_1..	18·0 ± 0·9 volts.

Azimuth 1.—This yielded the largest value of d, $d_1 = 9{\cdot}40$ Å, and consequently gave the smallest circle, which is shown in Pl. XVII. fig. 1. Measurements were made on the P-patterns for a number of plates, and the results are summarized in Table I. (*a*). These bands are indicated in the photograph showing the circle. The

angles given in the table were measured with reference to the central band.

Table I. (*b*) gives the values of Φ, calculated from the radii of the circles, along with the corresponding primary wave-lengths. For a wave-length of ·07 Å. and a camera length of 17 cm. the radius of this circle would be about 2·0 cm.

Azimuth 2.—Here the spacing d_2 was 6·64 Å. Pl. XVII. fig. 2 shows a typical photograph. The measurements on the P-lines and circles are given in Tables II. (*a*) and

TABLE II. (*a*).

Spacing calculated in Å.	Angle measured.	Indices assigned.	Calculated angle.	Spacing in Å, assuming D = 5·42 Å.
1·56	0°	222	0°	1·56
1·03	0°	333	0°	1·04
$0{\cdot}78_5$	0°	444	0°	0·78
$0{\cdot}90_5$	29°	531	29°	$0{\cdot}91_5$
1·59	58°	$13\bar{1}$	$58\frac{1}{2}$°	1·63

TABLE II. (*b*).

λ in Å.	Φ_2 (volts).	λ in Å.	Φ_2 (volts).
·0836	19·5	·0718	28·5
·0830	21·0	·0704	29·0
·0745	30·0	·0685	21·0
·0734	23·5	·0668	23·0
·0726	23·5		
·0725	23·0	Mean value of Φ_2	$24{\cdot}0 \pm 1{\cdot}1$ volts.

II. (*b*) respectively. For λ = ·07 Å. the radius of this circle was about 2·4 cm.

Azimuth 3.—In this particular case $d_3 = D = 5{\cdot}42$ Å. Pl. XVIII. fig. 3 shows the type of diffraction pattern obtained, while Tables III. (*a*) and III. (*b*) give the results of the P-line and circle measurements respectively. For λ = ·07 Å. the radius of this circle was about 2·7 cm.

Azimuth 4.—In this case d_4 was small, 3·83 Å., and the circle of large radius. Consequently it proved difficult to obtain much intensity in this circle without very much

over-exposing the remainder of the photograph. Shinohara *, in some work on P-patterns, has shown a photograph of the broad bands occurring for this azimuth, though the maximum diffraction angles obtainable in his case did not permit of any circle appearing. For $\lambda = \cdot 07$ Å. the radius of the circle observed here was about 3·2 cm. The results for this azimuth are given in Tables IV. (*a*) and IV. (*b*).

The diffraction patterns obtained with other settings of the crystal face were photographed, but, although

TABLE III. (*a*).

Spacing calculated in Å.	Angle measured.	Indices assigned.	Calculated angle.	Spacing in Å, assuming D=5·42 Å.
1·90	0°	202	0°	$1\cdot91_5$
0·97	0°	404	0°	0·96
$0\cdot64_5$	0°	606	0°	0·64
0·85	26°	602	26°	0·86
1·33	45°	400	45°	$1\cdot35_5$

TABLE III. (*b*).

λ in Å.	Φ_3 (volts).	λ in Å.	Φ_3 (volts).
·0834	17·0	·0727	16·5
·0808	12·0	·0722	19·5
·0765	15·0	·0714	25·5
·0740	14·5	·0703	13·0
·0738	15·5	·0673	12·0
·0728	18·0		
		Mean value of Φ_3..	$16\cdot5 \pm 1\cdot1$ volts.

broad central bands could be obtained, no other circles were ever seen. For these azimuths, however, simple considerations of the positions of the zinc and sulphur atoms showed that the intensity of these circles would be much smaller than that of any actually observed. Consequently any such circles were lost in the general background, and measurements of them were not possible.

Estimation of Φ.

In order to make any estimates of the mean potentials existing inside the tubes it is necessary to have models

* Sc. Pap. I. P. C. R. Tokio, xviii. p. 223 (1932).

of the zinc and sulphur ions. In actual fact * zinc-blende belongs to a type of crystal which is intermediate between the ionic and homopolar classes, but for our purpose we must consider the crystal composed of Zn^{++} and S^{--} ions. Moreover, for exact calculations the electron distributions about the nuclei would have to be known; but for an approximate calculation the number of electrons present in a thin shell of radius r will be expressed as a simple algebraic function, spherical symmetry being assumed for the distribution. With this model the electrostatic potential at any point inside

TABLE IV. (*a*).

Spacing calculated in Å.	Angle measured.	Indices assigned.	Calculated angle.	Spacing in Å, assuming D = 5·42 Å.
$1{\cdot}36_5$	0°	040	0°	$1{\cdot}35_5$
1·04	16°	151	16°	1·04
1·63	24°	131	25°	$1{\cdot}63_5$
$1{\cdot}11_5$	34°	242	35°	$1{\cdot}10_6$

TABLE IV. (*b*).

λ in Å.	Φ_4 (volts).	λ in Å.	Φ_4 (volts).
·0834	26·5	·0725	31·0
·0765	40·0	·0684	35·5
·0762	40·0	·0668	32·0
·0740	30·5		
		Mean value of Φ_4..	32·0±2·0 volts.

either ion, due only to the charges present in that ion can be calculated.

Many algebraic functions might be used as approximations of the electron distribution about the nucleus. We consider here an ion of nuclear charge +N, having a number N′ of extra-nuclear electrons. Then it is assumed that the charge density ρ at any point r inside the ion is given by

$$4\pi r^2\rho\, dr = \mathrm{K}(a-r)\, dr,$$

where K is a constant and a the ionic radius.

* W. L. Bragg, 'The Crystalline State,' vol. i. p. 128 (1933).

Since

$$\int_o^a \mathrm{K}(a-r)\,dr=\mathrm{N}'e, \quad \mathrm{K}=\frac{2\mathrm{N}'e}{a^2}.$$

In considering the potential at r, $\mathrm{V}(r)$ we find it in two parts, one due to the charge inside the shell of radius r and the other due to the charge outside. The two contributions are

$$\frac{e}{r}\left[\mathrm{N}-\frac{\mathrm{N}'}{a^2}(2ar-r^2)\right] \quad \text{and} \quad -\frac{2\mathrm{N}'e}{a^2}\left[a\log\frac{a}{r}-(a-r)\right];$$

so that

$$\mathrm{V}(r)=\frac{e\mathrm{N}}{r}-\frac{e\mathrm{N}'r}{a^2}-\frac{2\mathrm{N}'e}{a}\log\frac{a}{r}.$$

This expression has been evaluated for the zinc and sulphur ions. For Zn^{++}, $\mathrm{N}=\mathrm{N}_1=30$, $\mathrm{N}'=\mathrm{N}_1'=28$, and $a=a_1=0{\cdot}8$ Å. For S^{--}, $\mathrm{N}_2=16$, $\mathrm{N}_2'=18$, and $a_2=1{\cdot}8$ Å. There is some small divergence in the values of a_1 and a_2 given by Pauling and Goldschmidt *, so that intermediate values have been taken. In considering the space lattice we find that for every point lying outside the ions and having a positive potential, in consequence of the proximity of a zinc ion, there is another point having an equally negative potential because of the exactly similar proximity of a sulphur ion. For our purpose therefore we may take the potential at all such points to be zero: in some of the calculations this is perfectly correct, in others a good approximation. There is, however, an important exception: as the sulphur ion is larger than the zinc ion, there must be a region around the latter which has no exact opposite around the former. At a distance r from the zinc nucleus, where $a_2>r>a_1$, the potential will be

$$(\mathrm{N}_1-\mathrm{N}_1')e/r=2e/r.$$

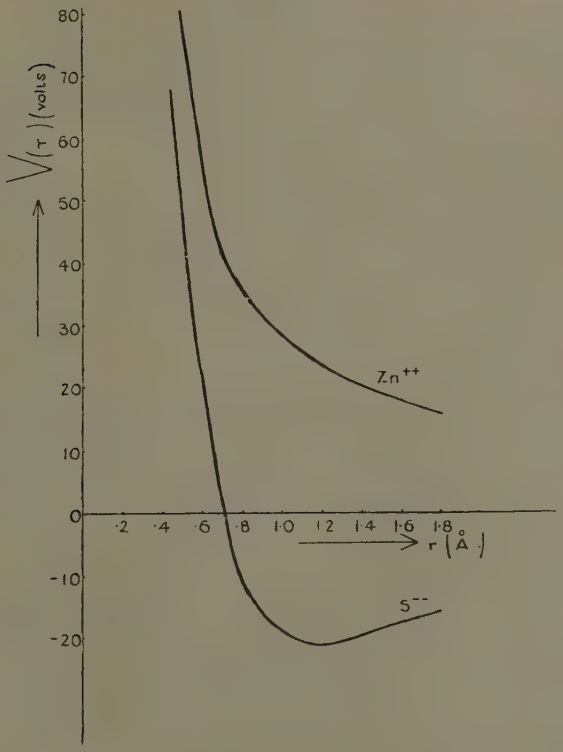
Fig. 3 shows the variation of V with r for both ions. The fairly constant negative value of $\mathrm{V}(r)$ for the sulphur ion when $r>1{\cdot}0$ Å. is important, as it simplifies the calculations which follow.

* *Handbuch der Physik*, xxiv. 2, p. 933 (1933).

Let us first calculate the mean inner potential Φ_0: the unit cube of side 5·42 Å. contains four zinc and four sulphur ions. Now

$$\Phi_0 \times 5{\cdot}42^3 \times 10^{-24} = \int V \, d\tau$$

Fig. 3.



taken throughout the cube, where $d\tau$ represents the element of volume.

$$\int V \, d\tau = 4\int_o^{a_1} V_1(r) 4\pi r^2 \, dr + 4\int_o^{a_2} V_2(r) 4\pi r^2 \, dr + 4\int_{a_1}^{a_2} 4\pi r^2 \frac{2e}{r} dr,$$

where the first integral, on the right-hand side of the equation, refers to the zinc ions, the second to the sulphur,

and the third to the space outside the zinc ions already discussed. Now it can be shown that

$$\int_0^a V(r)4\pi r^2\,dr = 2\pi e a^2\left(N - \frac{17}{18}N'\right),$$

and

$$\int_{a_1}^{a_2} 8\pi r e\,dr = 4\pi e(a_2^2 - a_1^2).$$

Thus we have, putting in all the appropriate values, $\Phi_0 = 9{\cdot}6$ volts. Using Goldschmidt's values of a_1 and a_2 we get a value of 9·3 volts, while Pauling's values give 9·6 volts. Here at least no serious inaccuracy arises because of the lack of accuracy in a_1 and a_2. The value found experimentally has been shown * to be 12·6 volts, so that the calculated value is smaller than the measured by some 20–25 per cent. Furthermore, this model of an ion has been used to calculate the mean inner potential of rock-salt, a true ionic crystal; the value obtained, 7·7 volts, agrees well with that found experimentally, 7·0–7·5 volts.

We will now consider the potentials inside the tubes shown in fig. 2, the diameter of any tube being taken to be that of the zinc ion, $2a_1$. This assumption appears at first sight to be quite arbitrary, but it must be remembered that, in addition to the fact that the zinc ion is heavier than the sulphur, the electron is much more likely to be guided into places of high electrostatic potential rather than into those of low (see fig. 3).

Azimuth 1.—We have to evaluate $\int V\,d\tau$ taken throughout the volume of the tube $\pi a_1^2 d_1$. This summation includes the contribution of one zinc ion, 205 volt Å.^{3}, that of a space outside the zinc ion but within the dotted circle (see fig. 2), 120 volt Å.^{3}, that of the core (of radius a_1) of the sulphur ion, 115 volt Å.^{3}, and that of a large outer portion of the same ion, -90 volt Å.^{3}. The contributions of small outer portions of other zinc and sulphur ions, which are seen from fig. 2 to be included within the tube, may be neglected, since for all practical purposes the potential in the very outer regions of the dotted circle around the zinc ion is exactly opposite to that of the outer parts of the sulphur ion (see fig. 3). Moreover,

* Phil. Mag. xviii. p. 656 (1934).

that portion of the tube, not included in any of the above volumes, may safely be ignored. Hence

$$\int V d\tau = 350 \text{ volt Å.}^3,$$

and this, in turn, equals $\Phi_1 \pi a_1^2 d_1$, Φ_1 here being our theoretical value.

Azimuth 2.—Proceeding in a similar manner, and taking into account those parts of zinc and sulphur ions cut off by this tube, we get a contribution of 325 volt Å.^{3} for the zinc ion and -30 for the two sulphur ions concerned. Hence

$$\int V d\tau = 295 \text{ volt Å.}^3 = \Phi_2 \pi a_1^2 d_2.$$

Azimuth 3.—Besides the usual portions of a zinc ion this tube contains the outer portions of two sulphur ions. Summing in the usual way, we get

$$\int V d\tau = 270 \text{ volt Å.}^3 = \Phi_3 \pi a_1^2 d_3.$$

But it can be shown that that portion of the tube not included in the ions already accounted for will on the whole have a negative potential, for part of it is occupied by sulphur ions in the next cleavage plane *. An accurate estimate of this contribution is difficult to make, but it is of the order of 30–40 volt Å.^{3} negative. Thus we may write with more accuracy

$$\Phi_3 \pi a_1^2 d_3 \sim 235 \text{ volt Å.}^3.$$

Azimuth 4.—Here again the contributions contain a very small one, due to ions in neighbouring planes, of the order of -10 volt Å.^{3}. The final value is

$$\int V d\tau = 255 \text{ volt Å.}^3 = \Phi_4 \pi a_1^2 d_4.$$

Discussion.

Table V. shows a comparison between the measured and estimated values of Φ_0 and $\Phi \pi a_1^2 d$.

These results show good numerical agreement, and seem to indicate that on the whole valid assumptions

* It is to be noted that for the two azimuths already considered the contributions of ions in the next planes may be ignored, as those due to zinc ions are cancelled by those due to sulphur ions; this present azimuth is exceptional in that more sulphur ions contribute than zinc.

have been made in the second half of the work. At the same time other models of the zinc and sulphur ions could be found to give equally good agreement. The present model was chosen for two reasons—firstly, it gave the right type of variation of ρ with r, and, secondly, it was simple to deal with mathematically. The assumption that ZnS is ionic was necessary in order to make any estimates of Φ whatsoever. As has already been pointed out, the actual composition of the lattice is not definite. Some of its properties indicate it to be ionic, others to be homopolar.

More important than the actual numerical agreement is the descending order of magnitude of $\pi a_1^2 d\Phi$ for the separate azimuths in the two rows of Table V. It is 1, 2, 4, 3 in both cases. This agreement justifies the

TABLE V.

	Φ_0 (volts).	$\pi a_1^2 d_1 \Phi_1$ (volt Å.²).	$\pi a_1^2 d_2 \Phi_2$ (volt Å.²).	$\pi a_1^2 d_3 \Phi_3$ (volt Å.²).	$\pi a_1^2 d_4 \Phi_4$ (volt Å.²).
Measured . .	12·6	340±17	320±15	180±12	246±15
Estimated .	9·6	350	295	235	255

procedure of taking the diameter of the tubes as that of the zinc ion; for to increase it to that of the sulphur ion, or to decrease it by any large extent, would result in the total loss of this order in the theoretical calculations besides entirely spoiling the actual numerical agreement. The assumption that the ions have spherical symmetry is not accurate, but any small distortions of the ions from the spherical shape will not affect the calculations made to any appreciable extent.

The second-order circle which can be obtained for stibnite has not been observed in the present work for any azimuth. In view of the smaller spacings used here this is not surprising, although it might have been expected that azimuth 1, where $d = 9{\cdot}40$ Å., would have given a faint second order. Here, however, another factor is at work. The S^{--} ions occur at one-quarter the way between the successive Zn^{++} ions. This is sufficient to decrease the second-order intensity by a factor of four;

the possibility of the circle being seen at all is thus much diminished. It is doubtful moreover whether any of the present conclusions would have been strengthened or new features introduced had any second-order circles appeared. Emslie has already shown that it is difficult to make accurate measurements on them.

We are now in a position to consider possible explanations of the circles in more detail. At first sight the circle in Pl. XVII. fig. 1 might seem to be the envelope of a family of P-lines, slightly distorted; but although Kikuchi lines do have their intensities and positions slightly affected when approaching one another, any such envelopes can be shown to be parabolas. Further, as Emslie showed, the radius of the second-order circle is $\sqrt{2}$ times that of the first, whereas the second-order envelope of P-lines has linear dimensions twice that of the first. The circles cannot therefore be explained in this way.

The results of the present investigation strengthen the view that the circles are the result of one-dimensional diffraction; but the localizing of the electron wave packet, as depicted by Emslie, is hardly a necessary part of the explanation. After suffering an inelastic collision with an atom the electron radiates as a spherical wave the intensity of which will not be the same in all directions. The probability of the electron continuing in a direction making only a very small angle with its original direction is very high. Since the row of atoms considered in this work has always made a small angle with the primary ray (fig. 1) the chances of an electron being directed along it are correspondingly quite large. Diffraction of the electron by this one-dimensional lattice can then take place. We must, however, consider the effect on the diffraction pattern of the atoms in neighbouring parallel rows. In distant parts of the crystal the spherical wave approximates very closely to a plane wave. This can be diffracted, giving rise in the pattern to a semicircle of spots, and not to a full continuous semicircle. But near the origin of the spherical wave a totally different state of affairs exist. It can be shown that the phases of wavelets scattered by the first fifty or so atoms in the row next to that containing the origin of the spherical wave do not form an arithmetic progression. These atoms cannot therefore influence the diffraction pattern,

the actual number of atoms so rendered incapable depending on the ratio of the distance between neighbouring rows to d. Similarly for the next row even more atoms are unable to give any sharp diffraction maxima. Therefore if we can assume that the diffraction occurs near the origin of the spherical wave the semicircle cannot be broken up. This assumption is quite feasible, as there is an extinction coefficient for electrons travelling in the surface of the crystal. Prof. Thomson * has calculated that the distance travelled by a 30 kv. electron in a crystal before suffering an inelastic collision is of the order of 4×10^{-6} cm.; this is the order of effect required in the present explanation. Emslie's idea of an almost complete localization is therefore unnecessary; moreover, it would lead to difficulties in the explanation of such electron diffraction phenomena as the formation of P-lines; but apart from this his explanation seems correct and complete.

In conclusion, I should like to thank Prof. G. P. Thomson for the kind interest he has shown in this work and for his helpful discussion on many points.

Summary.

The one-dimensional diffraction of fast electrons by the (110) face of zincblende has been investigated. It seems certain that the electrons often become confined to potential tubes in the surface of the crystal, which contain whole or parts of both zinc and sulphur ions. A simple theory is proposed which satisfactorily accounts for the measured values of the mean electrostatic potentials existing inside these tubes. Certain other important features of the measurements are also indicated by the theory. The diameter of the tubes is shown to be about that of a zinc ion, 1·6 Å.

The Imperial College of Science
and Technology, London.
November 1934.

* Phil. Mag. xviii. p. 640 (1934).

XLIII. *Experiments with Inverted Diodes having various Filament Cathodes.*

By J. S. McPetrie, *Ph.D., B.Sc., A.M.I.E.E.* (*From the National Physical Laboratory* *.)

Abstract.

The paper describes experiments made with a series of coaxial-electrode diodes in which the anode was the internal electrode. By using different types of cathodes it is shown that the velocity of emission of the electrons from the cathodes is not an important factor in determining the properties of these "inverted" diodes as electronic oscillators. It is deduced from this result that electronic oscillations in any valve must be due to the direct transit of electrons between the electrodes, and not to an oscillatory motion of individual electrons within the valve.

1. *Introduction.*

THE author described recently † a two-electrode valve suitable for the production of oscillations of high radio frequency. The electrodes in this valve were cylindrical, and the anode lay along the axis of the external cathode. The following simple hypothesis was put forward to explain the occurrence of the oscillations. In a cylindrical diode in which the cathode is external to the anode, any component of the velocity of emission of an electron not in a radial direction gives it an angular momentum about the anode, which remains constant as the electron approaches the anode. The transverse velocity of the electron increases, therefore, as its distance from the anode decreases, and if the anode diameter is sufficiently small the electron may escape capture by the anode. The direction of emission of the electrons from the cathode is more or less indiscriminate, so that by a suitable choice of electrode dimensions a large proportion of the electrons in such a valve may execute a to and fro motion about the anode. It is well known that electronic oscillations can be obtained from an ordinary

* Communicated by Dr. R. L. Smith-Rose.

† J. S. McPetrie, 'Wireless Engineer,' xi. p. 118 (1934).

cylindrical electrode triode when the grid is maintained at a high positive and the anode at a small positive or negative potential with respect to the filament. It has been suggested that the oscillations obtained in this case are due in some way to the to and fro motion of the electrons about the grid. The motion of the electrons in the inverted diode is similar to that in the triode with positive grid. If, therefore, the vibratory motion of the electrons in the triode produces oscillations of high radio frequency, similar oscillations should be obtainable from the inverted diode. In the paper referred to above, the author described experiments which proved that oscillations could readily be obtained from an inverted diode, and that the oscillations had characteristics similar to those produced by the triode with positive grid. The present paper describes further experiments which were carried out with inverted diodes of various types with a view to obtaining further knowledge as to their mode of operation.

2. *Experimental Procedure and Results.*

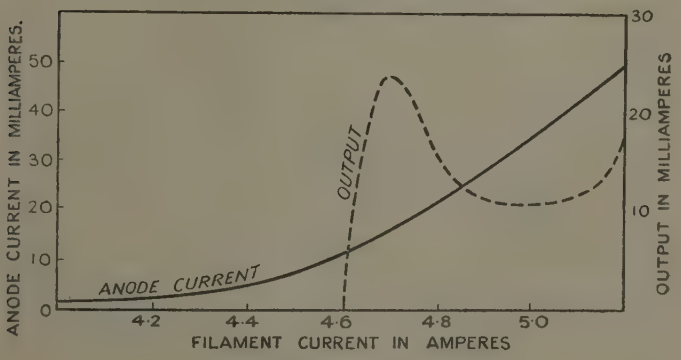
If the electronic oscillations in an inverted diode are due solely to the vibratory motion of the electrons about the anode, their intensity should increase with an increase in the number of electrons executing that motion. In the earlier paper an experiment is described in which a magnetic field was applied to an inverted diode in the direction of the electrode axis. The application of such a field should increase the number of electrons escaping capture by the anode, and, therefore, on the above hypothesis, should increase the oscillation intensity. It was found, however, that in all cases the output decreased when the magnetic field was used.

The proportion of electrons (in an inverted diode) escaping capture by the anode is a function of the velocity of emission of the electrons from the cathode. An experiment was made in which the output from an inverted diode was observed for a series of values of filament current, while the anode potential remained fixed at 600 volts. The solid-line curve in fig. 1 gives the relation obtained between the anode and filament currents, while the dotted line shows how the current in a Lecher-wire system coupled to the oscillator varied with the

filament current. Although the velocity of electron emission increases with filament current, it will be observed that the oscillation intensity does not follow the same law. Some factor other than the velocity of emission must enter, therefore, into the determination of the condition for maximum oscillation intensity.

The electric field in a cylindrical-electrode diode due to a difference of potential between the electrodes is greatest near the central electrode. In the usual arrangement for the production of electronic oscillations a three-electrode valve is used, and the grid is given a high positive potential with respect to the filament. An activated

Fig. 1.



filament could not be used under this condition, as the emissivity would be destroyed by the intense electric field at the filament. In the inverted diode, however, the field at the external cathode is small, so that a low-temperature cathode can be used without deleterious effect. A series of inverted diodes were constructed with cathodes in the form of a number of activated filaments in parallel in order to determine whether a high velocity of emission is required for the generation of electronic oscillations.

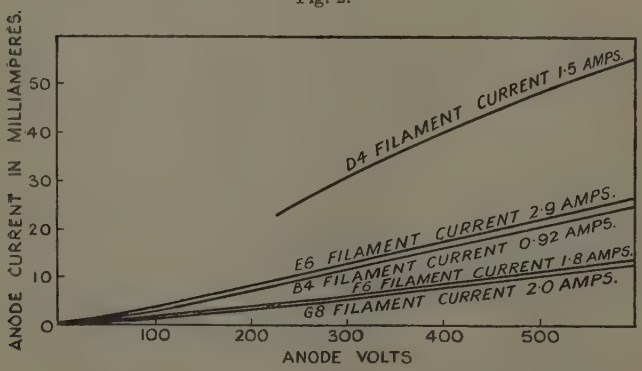
In the original bright-emitter diode described before, the cathode consisted of four tungsten filaments arranged in parallel between two rings, through the centres of which passed the anode. This method of construction was preserved in the new valves, but the number of

filaments in the latter varied from 4 to 8. The dimensions and other particulars of the valves are given in Table I., in which the original bright-emitter diode is also described.

TABLE I.

Marks.	Cathode.		Radius of cathode (cm.).	Radius of anode (cm.).	Radius of cathode / Radius of anode
	No. of filaments.	Type.			
A4.....	4	Pure tungsten.	0·8	0·064	12·5
B4.....	4	Thoriated tungsten.	0·4	0·019	21·0
C4	4	Coated or activated.	0·7	0·064	11·0
D4.....	4	,,	0·95	0·20	4·75
E6.....	6	,,	0·7	0·064	11·0
F6	6	,,	0·825	0·064	13·0
G8.....	8	,,	0·85	0·064	13·3

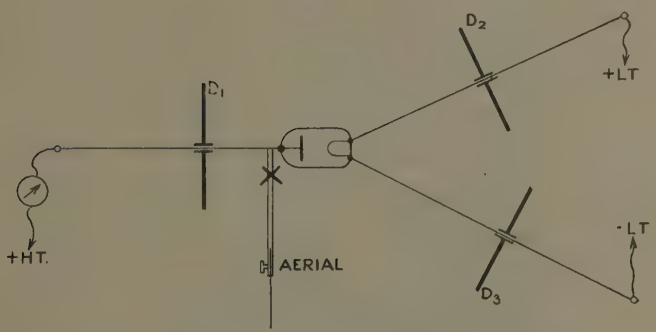
Fig. 2.



Static characteristics of anode current for typical values of filament current were taken for all valves over a wide range of anode potential. The actual filament current in each case was approximately that for which maximum oscillation intensity was obtained. The results are given in fig. 2, in which the valves are designated

according to the system given in the first column of Table I. The oscillatory properties of the valves were determined by connecting them in turn to the circuit shown diagrammatically in fig. 3. In this figure D_1, D_2, and D_3 represent copper disks, about 30 cm. in diameter, through the centres of which the battery-leads to the valve were passed. The disks introduce changes in the distributed constants of these leads, and when the valve is the source of oscillations stationary waves are produced on those parts of the leads between the valve and the disks. By a suitable choice in the positions of the disks resonance

Fig. 3.



Circuit used to determine dynamic characteristics of valves.

conditions of the leads may be obtained, and little oscillatory potential is found beyond the disks.

The oscillatory output was measured by the current in a thermojunction near the base of a $\lambda/4$ aerial attached to the anode terminal of the valve. For a given setting of the disks and a fixed anode potential the filament current was varied until maximum aerial current was obtained. The positions of the disks, the length of the aerial, and the filament current were then adjusted in turn, until the conditions were those for the maximum output for the given anode voltage. The tests were carried out for each valve for a range of anode voltages between 200 and 500, and the results of them are summarized in Table II.

The chief results to be noted from Table II. are that electronic oscillations can be produced by the diodes with low-temperature filaments, and that the outputs

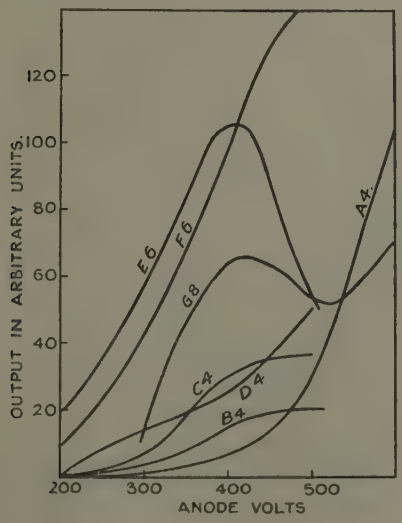
TABLE II

Valve.	Anode voltage.	Anode current (ma.).	Filament voltage.	Filament current (amp.).	Wave-length (metres).	Aerial current (ma.).
A4	400	20·0	6·5	6·0	—	1·75
	500	28·0	6·8	6·1	1·32	3·5
	600	29·0	5·3	5·2	1·24	7·8
B4	400	7·0	3·5	8·0	0·86	2·8
	500	10·0	3·8	8·5	0·82	3·5
C4	300	7·5	2·72	1·8	—	1·75
	400	12·5	2·27	1·5	1·32	4·0
	500	20·0	2·07	1·8	1·28	4·4
D4	300	30·0	2·1	1·5	—	2·7
	400	44·0	2·5	1·6	1·48	3·5
	500	45·0	2·1	1·5	1·38	5·25
E6	200	3·0	3·0	2·2	1·54	3·25
	300	7·0	3·6	2·6	1·28	5·7
	400	15·2	4·25	2·9	1·16	7·6
	500	34·0	5·0	3·2	1·06	5·25
	600	34·0	4·5	3·0	0·96	6·25
F6	200	2·5	1·25	1·5	1·9	2·5
	300	7·0	1·8	2·0	1·62	4·6
	400	7·5	1·75	1·8	1·3	7·5
	500	20·0	2·2	2·2	1·2	10
G8	300	7·0	2·0	2·6	—	2·6
	400	6·5	1·35	2·0	1·37	6·0
	500	12·5	1·75	2·3	1·3	5·4

from them are greater than that from the somewhat similar diode having a high-temperature cathode. The latter point is brought out more clearly in fig. 4, in which the outputs from the various valves are shown plotted for a range of anode voltages. It will also be seen from

this figure that the outputs from the diodes having cathodes in the form of 6 and 8 filaments were greater than those obtained with the 4-filament cathode valves. The energy dissipated in the cathode of any valve increases with an increase in the number of filaments forming the cathode. The overall efficiency, as well as the output, however, appears to be greater for the valves having the greatest number of filaments. The results given in Table II. have been used to obtain the curves in fig. 5,

Fig. 4.

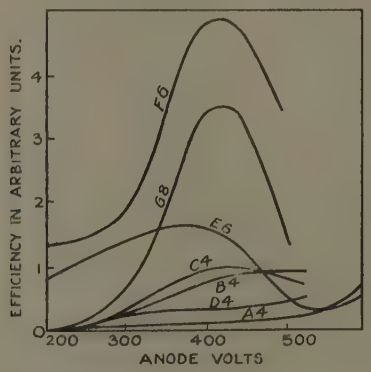


which show the efficiency in arbitrary units for the various valves for the different anode voltages. The general shapes of the curves in this figure are very similar to those in fig. 4, but the dull-emitter diodes are shown to greater advantage when efficiency is considered. The 8-filament diode was inferior to the two 6-filament valves in output and to one of them in efficiency. This may be explained by the fact that when current was passed through the filaments in the 8-filament valve they were displaced appreciably, due to the electrical interaction between them.

3. *Discussion of Results.*

The main conclusions which may be drawn from the experiments described above are: (*a*) electronic oscillations can be obtained in an inverted diode when the velocity of emission of the electrons from the cathode is small: (*b*) the output and overall efficiency of the inverted diode increases, in general, with increase in the number of filaments forming the cathode. This result suggests that maximum output would be obtained by using an indirectly heated cathode.

Fig. 5.



The first conclusion is of some importance in the study of electronic oscillations, and its implication will be described in some detail. As outlined above, the to and fro motion of the electrons about the positively charged grid of a triode is supposed in some way to give rise to electronic oscillations. There is little doubt but that the conditions required for producing such oscillations in the triode are similar to those necessary in the case of the inverted diode. This means that, if it can be shown that in one type of valve the production of electronic oscillations is not primarily dependent on a vibratory motion of the electrons within the valve, then such a motion is not necessary in the other type of valve.

It was stated above that the application of an axial magnetic field to an inverted diode decreased the output

from the valve, although the effect of such a field should be to increase the number of electrons vibrating about the anode. The results given graphically in fig. 1 show that for a given anode potential there is a definite value of filament current for the condition of maximum output. As pointed out above, this fact proves that increase in filament current and, therefore, in the velocity of emission of the electrons from the cathode, does not necessarily increase the output from an inverted diode, although it certainly increases the proportion of electrons executing a to and fro motion about the anode. Finally, the experiments described immediately above prove conclusively that a high velocity of electron emission is not a requirement for strong oscillations; in fact, the results given in fig. 4 show that the output from a bright-emitter diode is appreciably less than that from a valve of similar construction, but with activated filaments operating at a low temperature.

The primary reason for the oscillations obtained from an inverted diode cannot then be due to the vibratory motion of the electrons within the valve. The wave-length of the oscillations in both the inverted diode and the triode with positive grid is approximately proportional to the time of transit of the electrons across the valve. In the inverted diode, therefore, the oscillations are due probably to a condition of negative resistance being obtained over a narrow band of frequency, the position of which is determined by the time of transit of the electrons from the cathode to the anode, and not, as originally suggested, due to the oscillatory motion of individual electrons within the valve. This alternative hypothesis would explain the experimental result that the inverted diodes with low-temperature filaments were found to give stronger oscillations than those valves with pure tungsten filaments, as the proportion of electrons captured by the anode on the first transit across the valve would be much greater in the low-temperature cathode valves.

The experiments with the inverted diodes constitute the first direct proof that oscillations of high frequency can be obtained from a diode when electrons pass directly from the cathode to the anode. As a to and fro motion of the electrons appears to be unnecessary in the case of the inverted diode for the production of electronic

oscillations, it may be argued that the usual hypothesis of a to and fro motion about the grid of a three-electrode valve when used for the same purpose is not correct. The negative resistance condition within the triode must be explained in a manner similar to that in the inverted diode, namely, the direct transit of electrons from one electrode to another. The triode with positive grid can be considered as two diodes connected together, with the grid as their common anode. This can readily be seen when the motion of the electrons within the valve is considered. The electrons leaving the filament are attracted by the positive grid, which they reach with a high velocity. A certain proportion of electrons escape into the grid-anode space. These electrons are brought to rest a small distance in front of the anode, as it has a small negative potential with respect to the filament. From this zero-velocity position the electrons are attracted back to the grid, which they reach with the same velocity as they had when passing it originally in the opposite direction. The zero-velocity position in the grid-anode space may be considered as a cathode emitting electrons towards the grid. The conditions within the triode can be seen in this way to approximate to those of two diodes having the grid as their common anode. One of these diodes has its filament within the anode, while the other is similar to the inverted diode in having the anode within its cathode. The experiments described above show that electronic oscillations can be obtained in an inverted diode. The reason for electronic oscillations occurring in a triode with positive grid may be due, therefore, to the distribution of potential within such a valve forming an inverted diode within the grid-anode space.

No attempt has been made to explain why a condition of negative resistance should be possible within an inverted diode. Benham * has shown analytically that oscillations of high frequency may be possible within diodes having parallel-plate electrodes.

4. *Acknowledgments.*

The work described in this paper was carried out as part of the programme of the Radio Research Board,

* W. E. Benham, Phil. Mag. xi. p. 457 (1931).

and acknowledgment is due to the Department of Scientific and Industrial Research for granting permission to publish. The author is also indebted to Dr. R. L. Smith-Rose for advice on the method of presentation of the paper, and to Mr. H. M. Bristow, B.Sc., for assistance with the experiments.

XLIV. *Notices respecting New Books.*

Faraday. By Thomas Martin. [Pp. 143.] (Duckworth. Price 2*s.* net.)

THIS sprightly little volume is the fortieth in the "Great Lives" series of Messrs. Duckworth. Looking through the list of the forty or so volumes published to date one cannot but be struck by the fact that about half are writers—prose or poetry,—half a dozen are musicians, and the rest in descending numerical order are monarchs, scientists, politicians, divines, and a cricketer. The fact is, of course, that usually the lives of great scientists are generally too far removed in their sphere of action from general experience—it is only when their influence on life becomes a recognised part of it that a biography becomes a publishing possibility.

Everyone now knows a little about Faraday—at any rate every schoolboy does,—and this book puts further knowledge of the man himself within the reach of everyone whose appetite may perhaps have been whetted by the recent celebrations in London.

Relativity, Thermodynamics and Cosmology. By R. C. Tolman. [Pp. 1–501.] (International Series of Monographs on Physics.) (Oxford: Clarendon Press, 1934. Price 30*s.*)

To-day the theory of relativity has ceased to be the study of the select few and has become a part of the recognised training of most mathematical physicists. The need for a treatise which would start from the elements and work up to the advanced portions of the subject had become more and more acute in recent years. There can be no doubt that Professor Tolman's book will admirably fill this gap. His exposition of the special theory of relativity and its consequences in mechanics, electrodynamics, and thermodynamics is particularly good and most opportune. The presentation of general relativity is also excellent, and more special solutions of Einstein's equations and useful formulæ are given than is usually the case in works on relativity.

Nearly a third of the book is devoted to the cosmological applications of general relativity, which are exhaustively covered both from the observational and the theoretical points of view. (Milne's kinematical theory has, of course, to be excluded owing to its different line of approach.) Professor Tolman's attitude towards these cosmological problems will have a sobering influence on the over-confident. He insists that the various types of "expanding universe" are not alternative schemes according to one of which the "real" universe must be constructed. They are merely attempts to show that the phenomenon of the recession of the nebulæ can be brought within the scope of a theory which has already proved itself a satisfactory generalization of Newtonian mechanics.

An extension of thermodynamics to general relativity proposed by the author receives a good deal of attention. It is difficult, however, to see what observational facts this theory covers. It remains as yet, like the unified field theories of electricity and gravitation, largely formal in character.

A word should be said about the excellent way in which the various subjects—special theory, general theory, mechanics, electrodynamics, and thermodynamics, are presented in separate chapters. A beginner would have little difficulty in studying any one of them separately. Yet the book as a whole forms an invaluable and complete work of reference for the expert in the field of relativity and its cosmological applications.

A Mathematical Treatise on Vibrations in Railway Bridges. By Prof. C. E. Inglis. [Pp. xxv + 203.] (Cambridge University Press, 1934. Price 21*s.* net.)

The author was a member of the Bridge Stress Committee established under the Chairmanship of Sir Alfred Ewing in 1923, which published its report in 1928. The main objective of this treatise is the analysis of the oscillations due to hammer-blows and the production of formulæ for computing dynamic deflexions and the bending-moments and shearing-forces induced thereby.

The main causes of impact for which the Committee decided that allowance should be made have been enumerated, viz. (1) engine hammer-blows, (2) lurching of engine, and (3) rail-joints.

Only the first lends itself to mathematical investigation, and in this book Professor Inglis has investigated the vibration of a bridge due to hammer-blow. Impact effects due to rail-joints are quite incalculable and must be allowed for by empirical formulæ. The "Pencoyd" and other such empirical

formulæ do not take into account the main cause of the impact, namely, the action of the locomotive's balance weights. In fact, these impact effects depend on factors which are very different from any of those which form the basis of the older empirical formulæ.

Professor Inglis found it possible, with very considerable mathematical labour, to deduce a formula for the vibration of a bridge in any circumstances, and to obtain the complete diagram of oscillation of certain bridges, which had been tested at a range of speed, so as to produce the effects actually observed in testing. The analysis is based upon the fundamental principle that the deflexion of a bridge depends sensibly upon the primary harmonic component of the applied load. The results of these calculations agree remarkably well with those recorded.

The author deals first with the more simple analyses, and later considers the behaviour of bridges under moving loads coupled with hammer-blows. The problem is shown to be fundamentally the same as that of a forced oscillation on a loaded spring. The work is then pursued by means of harmonic analyses, and formulæ of a comparatively simple character have been evolved which yield results in very close agreement with those based upon more elaborate mathematics. An evaluation of the maximum oscillation set up by the passing of any particular locomotive has been obtained by substituting what he has termed a skidding locomotive for the travelling one at the centre of the bridge. The errors introduced by this ingenious method actually amount to a very slight over-statement of the stresses.

The important phenomenon of resonance is discussed, together with the effects of damping, length of bridge, and other characteristics. Attention is called to the fact that it is undesirable to run a locomotive at slow speeds over long-span bridges which would result in resonance and in the locomotive behaving as a springless mass. In the case of medium and short-span bridges the modifying influences of friction in the locomotive springs are deduced, the importance of damping considered, together with the fact that the riding of some engines at speed would cause violent oscillations in the springs due to the bridge itself.

The author is to be congratulated on the able mathematical investigation he has made and for the noticeable service he has rendered to engineers by his continued efforts to solve some of the problems of impact so as to obtain formulæ suitable for the drawing-office. The position in which bridge engineers find themselves to-day is considerably improved by the fact that we have available for our use a method by which we can calculate the effects of what seems to be the

most serious cause of impact. This work has enabled us to gauge the relative importance of the various contributing factors producing oscillations and disturbances in the bridge.

It is important, however, to indicate that these impact contributing factors will not assume the same relative importance in Colonial and Continental railways as in this country, and that it may not be appropriate to deal with the situation on the same lines as have been found to apply to the special conditions in this country. The scientific method propounded by the author is based upon conditions where hammer-blow is the most important factor. It ignores the effects due to the lurching of the locomotive and the rail-joints, which may be of primary importance in the case of short-span bridges. When locomotive engineers have reduced considerably the magnitude of the engine's hammer-blows, the loads prescribed by the Bridge Stress Committee will need revision. The task of carrying out such a revision will be greatly reduced by these simple formulæ evolved by the author. Further research will afterwards be necessary with the object of ascertaining the important bridge and locomotive characteristics, viz., the periodicity of oscillation of the springs, the frictional damping of the oscillation.

It is only possible to tabulate the necessary impact-allowances when all bridges and locomotives conform to some standard types. Until such times impact formulæ must be expressed in terms of the bridge and locomotive characteristics which decide the type of oscillation.

In any railway system the required characteristics will have to be determined by full-scale experiments.

Resonance Radiation and Excited Atoms. By A. C. G. Mitchell and M. W. Zemansky. [Pp. xvi + 338.] (Cambridge Series of Physical Chemistry.) (Cambridge: University Press, 1934. Price 18*s*.)

The purpose of this book is to give a comprehensive account of experimental and theoretical work on resonance radiation and related phenomena. Many of the topics are among those to which the classical electron theory was applied. In spite of its quantitative failure, its qualitative success, combined with the formal elegance of the treatment, resulted in the persistence of classical modes of expression until well into the quantum era. Moreover, in the first phase of the quantum theory few phenomena were more awkward than those connected with dispersion, scattering, and absorption. It is only with the later development of quantum mechanics that a satisfactory self-consistent method of theoretical attack on the problems involved became possible. The time has

been well chosen for a general review of the very considerable amount of experimental work which has been carried out in this field.

In an introductory chapter the fundamental ideas on line spectra and excited states of atoms are briefly discussed, and suitable sources for exciting resonance radiation and resonance lamps are described. In the next chapter stepwise radiation, sensitized fluorescence, and sensitized chemical reactions are dealt with. The theory of absorption lines is then very fully discussed, the various methods for determining the mean lives of excited atoms are considered, the experimental results being summarized in a useful series of tables. A systematic account of collision processes involving excited atoms is then given, and the theory of broadening of lines due to natural damping, the Doppler effect, and pressure effects is developed. The final chapter deals with the polarization of resonance radiation, with an account of the effects of magnetic and electric fields.

The book may perhaps be described as a classified series of accounts of experimental and theoretical work; as such it is singularly complete, and the accounts are extraordinarily clear and straightforward. The book as a whole, however, seems to suffer from the defects of the method adopted, particularly on the theoretical side. The theory appears in disjointed fragments, as required for the experiments under discussion; and it is difficult to gain a clear impression of the main underlying ideas. This drawback will affect the general reader more than the specialist, who will require particular pieces of information, which he will readily be able to find from the detailed table of contents.

That the book is included in a series on physical chemistry suggests a rather catholic interpretation of the term "chemistry," but possibly physicists are now presumed to be exclusively concerned with atomic nuclei. The book is well produced, but, from such a press, the drafting of the descriptive paragraph on the jacket is so careless as to call for surprised comment.

For those engaged in research (whether as chemists or physicists) in the field with which it deals this book will be indispensable: for the non-specialist it fills a very glaring gap in the recent literature of physical optics.

Modern Acoustics. By A. H. Davis. [Pp. 345 + xi.] (Bell & Sons, Ltd. Price 26*s*.)

The aim of this book, as set out in the Preface, is "to give the essentials of acoustics as the subject is practiced to-day"; the book is intended to be "of value to students, to technical

and research workers in acoustics, and to those in trade and industry who need to acquaint themselves with modern acoustical methods and to apply them."

The book is essentially a practical one and—as the author states—can be read by anyone who will take mathematical equations for granted. It is a difficult matter, however, to divorce a practical discussion completely from the related mathematical theory without making the discussion superficial, and it is specially difficult in certain sections of modern acoustics which involve carrying over the mathematical methods and terminology of the theory of electrical oscillations to the theory of acoustic vibrations. Dr. Davis has attempted this difficult task with considerable success; his treatment is certainly not superficial. But in certain parts the discussion is so concise that most students will probably find some difficulty in following the development of the argument. For example, in discussing the equations for the growth and decay of sound in rooms, the reviewer thinks that a simple proof of these equations should have been given; the thoughtful student will probably be puzzled when he reads on p. 149 that the mean free path for a sound ray in a closed room is $4V/S$, as shown by a statistical method. A reference to a very useful *résumé* by Eckhardt on this subject in the 'Journal of the Franklin Institute' might well have been given.

On the practical side the book is exceptionally good. The sections on the measurement of sound intensity, on sound analysis, and on noise are specially interesting. The chapter on the acoustics of buildings is somewhat brief, but briefness here is justified by the earlier book of Dr. Davis and Dr. G. W. C. Kaye and by the appearance of several other books on this subject during recent years. Another chapter discusses the reproduction and recording of sound; the general principles are described, and technical details are for the most part omitted. Recent work on supersonic waves is mentioned only briefly in connexion with piezo-electric sources of sound.

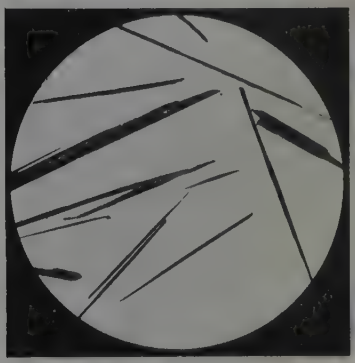
The general appearance of the book is very pleasing. It has many good diagrams and a number of interesting plates. A good index to authors and subjects is given, and references to original papers are given at the bottoms of the pages. The book will be very useful to technical and research workers, and will provide an excellent introduction to modern acoustics for those in trade and industry; for students it will prove an interesting supplement to more prosaic lectures and books.

[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]

Initiation of catalytic combustion on the surface of a platinum wire in CO and air, at a temperature of about 1360° C.

Fig. 1.

Fig. 2.



Picryl aniline.

Prisms. Needles.

Fig. 5.

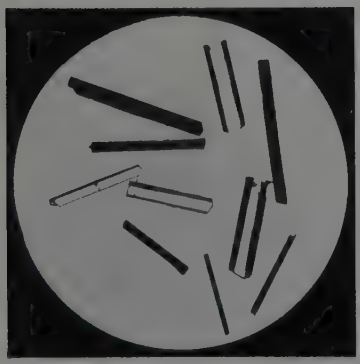
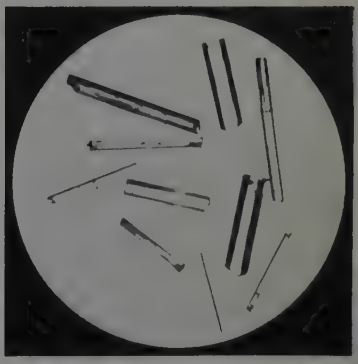


Fig. 6.



Picryl *p*-toluidine.

Both forms.
(Yellowish-green filter.)

Both forms.
(Red filter.)

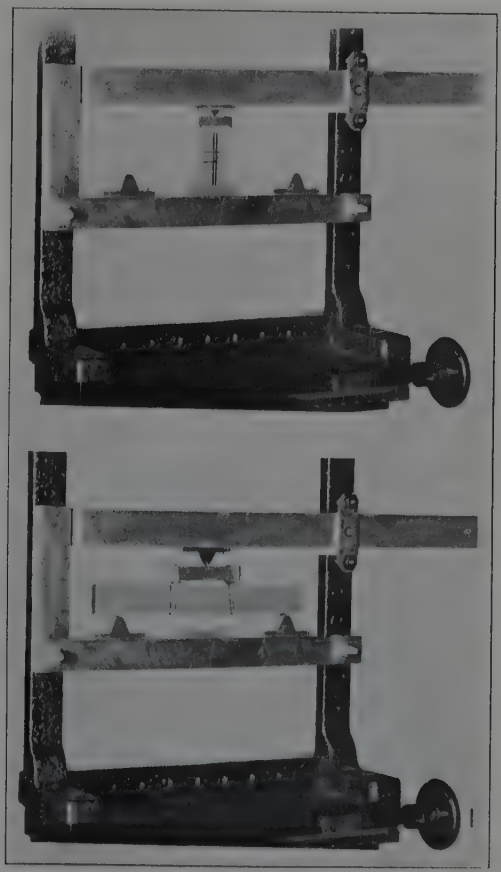
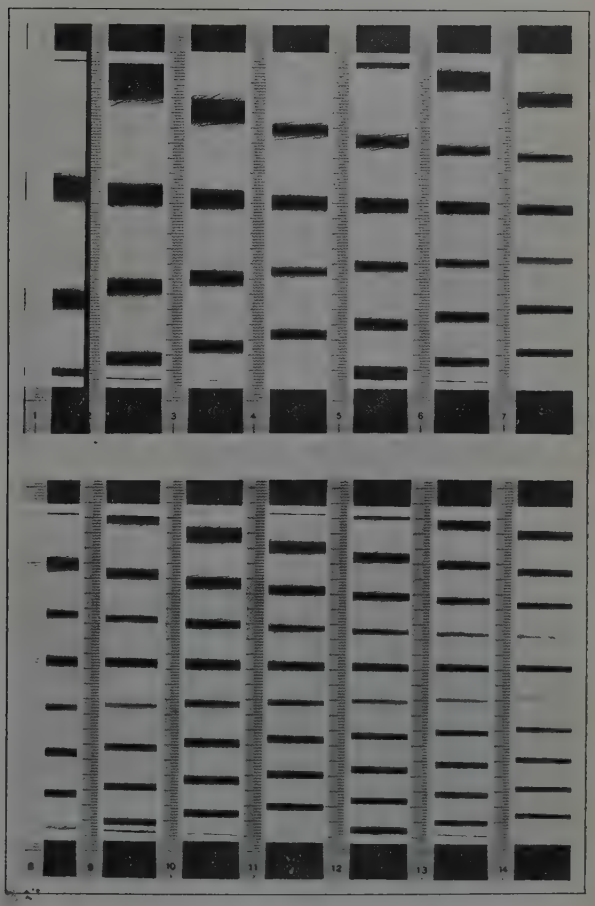


Fig. 2 (upper).—The compression specimen.
Fig. 4 (lower).—The bending specimen.

FIG. 5.



Isochromatic lines in the beam.

FIG. 1.

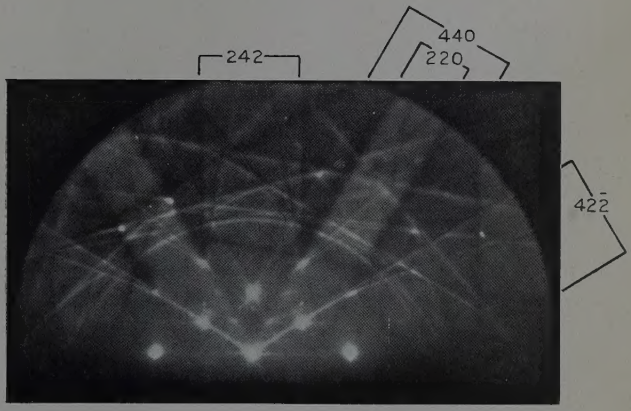


FIG. 2.

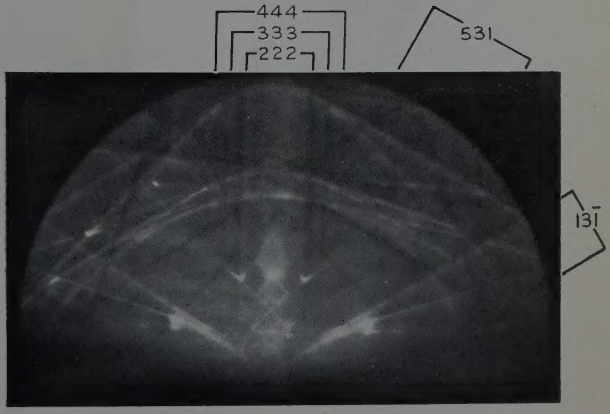


FIG. 3.